
Infectious Diseases, Public Policy, and the Marriage of Economics and Epidemiology

Mark Gersovitz • Jeffrey S. Hammer

The assumption of rational choice helps in understanding how people respond to infectious diseases. People maximize their well-being by choosing levels of prevention and therapy subject to the constraints they face. Objectives and constraints are numerous, necessitating tradeoffs. For example, this approach predicts how people respond to changes in the risk of infection and to the availability of diagnostic tests. The combination of individual rationality with epidemiological models of infection dynamics predicts whether individual choices about infectious disease prevention and therapies produce the best possible social outcomes. If not, individuals' choices generate rationales for government interventions to influence the levels of preventive and therapeutic activities. Optimal policy usually means accepting endemic infection, but at a level lowered by a coordinated package of interventions. Economics combined with epidemiology provides much qualitative guidance on the design of such packages, including immunization programs.

The recent marriage of economics and epidemiology, though long delayed, promises to be fruitful. Public health means public policy, never more so than with respect to infectious diseases. Economics has much to say about public policy, and epidemiology has much to say about the dynamics of infectious diseases. This article reports on the first progeny of this marriage—vigorous and precocious in what they say about public policy, but with potential still to fulfill.

Infectious disease has two modes of transmission: from person to person, as with HIV/AIDS, or by vectors, such as mosquitoes, as with malaria or yellow fever. Both modes are strongly influenced by how people choose to act both individually and collectively through their governments. People make choices about prevention and therapy from among the options available to them. An important influence on these

choices is the probability of becoming infected, which varies with the proportion of people who are infectious, among other factors. Choices about prevention and therapy in turn affect the aggregate infection rate, producing a system of simultaneous feedback. Classic models of infectious disease (Anderson and May 1991) have not, however, incorporated such endogenous behavioral mechanisms into their dynamics.

One way to do so is to postulate that individuals make rational choices, the traditional assumption of economics. Once individual behavior is introduced into the analysis, a natural question is whether individual behavior is the most desirable based on social criteria or the aggregation of the costs and benefits for all people of each individual's behavior. Answering this question is the bread and butter of economics, essentially assessing the validity in the particular circumstances of the health sector of Adam Smith's hypothesis of the Invisible Hand.

When individuals do not take into account the effects of their actions on others, either negative or positive, there are externalities and a justification for government intervention to align the private costs and benefits with the social costs and benefits of decisions. A major theme in the economics of public health is how and why these externalities occur and what can be done about them.

A number of economists have taken up the task of applying economics to public policy toward infectious diseases, beginning with the foundation block of individual rational behavior and working up to aggregate dynamics and the design of policy. An important concept is the optimal policy package, the optimal rules for coordinating preventive and therapeutic expenditures over the course of an epidemic. Although doctors and patients naturally group actions into prevention (avoiding infection) and therapy (mitigating infection), this distinction is not a good guide to the design of an economically efficient policy package.

On some economic criteria and for some infections, there is no distinction between prevention and therapy. For instance, when optimally chosen the two types of inputs may move together over the course of an epidemic, increasing or decreasing in step. Such synchronization need not occur, however, and reliance on prevention may decline as reliance on therapies rises or vice versa. The targeting of inputs is important in determining this pattern, for instance, whether treatment is directed at everyone or just the infected.

Both prevention and therapy likely involve externalities. In this regard, the important influence on the design of the policy package is whether the discrepancy between individual and social incentives to prevent disease is the same as that between individual and social incentives to adopt therapies. These discrepancies are equal for diseases from which people recover only to become susceptible again, as with curable sexually transmitted diseases, such as chlamydia and gonorrhea. In these cases it is optimal for government to subsidize preventive and therapeutic activities equally. For other diseases, the government may need to provide quite different financial incentives for people to undertake preventions and therapies. The

situation of vector-borne diseases is even more complicated. These diseases typically respond to many preventive actions. Some preventive actions share economic attributes of therapies, and others differ both from therapies and from other preventive actions.

Many of the examples of decisions about infectious diseases in this article come from the experience of poor countries and, in particular, from the experience with HIV/AIDS. First, infectious diseases are much more prevalent in poor countries both because their geoclimatic environments are more conducive to many of these diseases, especially those involving vectors and macroparasites, and because individuals and governments have taken many fewer of the actions that control infectious diseases in richer countries. Second, much of the work by economists on infectious diseases has been on HIV/AIDS in both rich and poor countries. HIV/AIDS, a threat everywhere, has become the most important health concern in many of the heavily infected countries of Africa. But the article also discusses the economics of diseases from which people recover and diseases spread by vectors and explores options for dealing with diseases, such as vaccines and curative medicines, which are unfortunately not available for the control of HIV.

Constraints, Objectives, and Strategic Behavior in Individual Choice

The rationality postulate supposes that individuals make the best choices, as they see them, from the options available to them subject to the information they have. As with all applications of this postulate, two things need to be specified for it to be useful in making predictions. First are the constraints that the individual faces: the options for action as determined by the disease environment, the information the individual has about the environment, and the costs of being sick and dealing with the disease through prevention and therapies. Second are the objectives of the individual, or how the individual ranks different choices among these options.

Constraints

The biology of a disease sets many of the constraints. Once infected, does an individual recover and become immune or susceptible again or suffer chronic ill health or death? How does the individual value these circumstances? The answers to these questions determine part of the costs of being infected. Can individuals know when they are infected and infectious, and when others are infectious, whether through symptoms or a medical test? Can therapy alter the course of the disease? If so, the costs of the disease include the costs of illness as mitigated by therapy plus the costs of the type of therapy chosen.

Now moving backward to the point at which a susceptible person is exposed, is the disease transmitted easily or with difficulty and how? What preventive actions can a person take to lessen the risk of exposure to infection, and how much do these preventive actions cost? Is a vaccine available? Is a vector such as a mosquito involved, and how can it be combated? To what extent is the risk of infection determined by the overall rate of infection and hence the stage of the epidemic? Are the infectious people who put a susceptible individual at risk anonymous, as in the case of influenza, or plausibly identifiable to the susceptible, as with a sexually transmitted disease?

Information also constrains how individuals respond to infectious diseases. Two types of information are relevant to individual decisions about preventive and therapeutic actions. The first type is general information, including information about the existence of a disease, how it is and is not transmitted, whether infectious individuals can be asymptomatic, and whether preventive, palliative, and curative options are available and at what cost. An important question is whether such general information is the real constraint on people's decisions. For a newly emerging disease, such as HIV/AIDS in the 1980s, gaining general information through research and disseminating it to the population as a whole are critical. But after the initial phase there may be much less scope for affecting behavior through general information, an important exception being young people if parents and other adults do not effectively communicate knowledge. If people already have this type of information, providing it over and over is unlikely to change how people behave. If individuals make rational choices, then to be effective interventions need to affect individuals' constraints.

The second type of information is specific to the individual. It includes information about whether the individual is infected or infectious, whether the people with whom the person consorts are infectious, and the risks in the immediate environment, such as disease vectors or the presence of macroparasites in soil and water. There may be no way to tell without a test whether someone is infectious. Of course, people may have useful information about whether they are themselves more or less likely to be infected and, especially in the case of sexually transmitted diseases, about which prospective contacts are more or less likely to be infected. Such information is better than nothing—sometimes much better—and can be used to condition choices. Better yet is a medical test. If a medical test exists, the decision to be tested raises many strategic questions for individuals that fit well within the rationality paradigm, as explored later in this article.

Objectives and Tradeoffs

A person's decisions affecting health are influenced by objectives as well as constraints. Though reflecting simple common sense, these objectives are often neglected in the formulation of health policies. That oversight is important because

knowledge of people's objectives helps predict their behavior and because most economists believe in consumer sovereignty, which means taking into account how people value their own circumstances. Individuals want to minimize the money costs of prevention and curative or other therapeutic activities insofar as they pay these costs. In addition, all the costs of prevention, including all the things that people forgo to avoid being infected, are also costs of a disease just as much as the pain, fear, loss of income and other opportunities, early death, and other direct costs incurred as a result of becoming infected.

People value many things in addition to health. They value physical intimacy with other human beings, including sexual relations, whether for pleasure or procreation. But physical intimacy promotes the spread of infection, whether sexually transmitted diseases or diseases of proximity such as influenza, tuberculosis, and leprosy. To the extent that individuals are altruistic, they also value the costs paid by people whom they might infect, such as their sexual partners. People may also want their sexual partners to believe that they are sexually faithful and they may want their partners to be sexually faithful independently of the risk of infection. People also value moving around freely, even in environments that expose them to the risk of infection by vector-borne and macroparasitic diseases. They prefer not to boil water, sleep under bednets, or restrict their outdoor activities. They want the benefits of going into forests even at the risk of malaria, yellow fever, and plague or of working on their farm even at the risk of schistosomiasis (bilharzia) or onchocerciasis.

Thus, many goals are in conflict, requiring tradeoffs. A basic tradeoff is between expenditures to augment health through prevention or therapy and expenditures on other goods and services that augment well-being in other ways. As people sacrifice successive units of expenditures on things other than health to increase their expenditure on health, their valuation of these successive sacrifices tends to rise, sometimes sharply. This rising marginal utility of what is sacrificed is a natural brake on how many resources people want to divert from other expenditures to health. Most people do not want to dedicate all their resources to improving their health.

All these considerations are irrelevant if individuals lack choices. For infectious diseases people can almost always pursue various preventive actions and they may have the option of therapies. For many diseases there is a continuum of increasing effort for prevention and therapy. People can wash their hands or boil water thoroughly, more thoroughly, even more thoroughly, and so on. Similarly, people can pursue and comply with therapeutic regimens to varying degrees. In many cases the effectiveness of successive equal increases in health inputs tends to diminish. This diminishing marginal product of inputs is another brake on how far people are prepared to increase health expenditures, complementing the increasing valuation of additional forgone units of expenditure on things other than health.

Rational Fatalism

An important determinant of choices about prevention is the infection rate among potential partners, casual contacts, and vectors, which in turn is likely to depend on the overall infection rate. In the case of AIDS, an incurable disease, models of rational choice have found that the number of partners a person chooses may not be negatively related to the infection rate of their potential partners, a sort of rational fatalism.

To illustrate this point, assume that people do not care about infecting their partners. For ease of argument, assume further that people get great utility from one (randomly chosen) partner and declining but positive utility from subsequent ones. All these partners are equally likely to be infected (and infectious). At moderate levels of overall HIV infection, people choose to have more than one partner but less than the number they would choose in the absence of the infection. By assumption, if the infection probability of all partners rises substantially, people do not cut back to less than one partner. Because their first partner is likely to infect them, more than one partner poses little additional risk. People no longer cut back the number of partners after the first and might even increase the number of partners back to the maximum they would have had were there no disease at all.

This conclusion is, of course, the outcome of a specially chosen example. Under different assumptions, such as an infection rate rising from negligible to moderate levels, people might react in the opposite way, cutting back on the number of partners. The general point is that rational people who weigh both a desire for more partners and the probability of infection may choose to increase the number of partners in response to an increase in the probability that any one of their potential partners is infected. Thus there is a rational explanation for fatalistic and seemingly irrational behavior (Philipson and Posner 1993, p. 49; Kremer 1996).

The notion of rational fatalism is relevant to choices about the intensity of all sorts of exposure, not just to HIV/AIDS. For example, each day that a person extends a trip into malarial forests might be analogous to an additional partner, presuming that after having been infected one can continue to be active, at least while incubating the disease. Rational fatalism will be less likely if a pathogen exhibits superinfectivity, that is, when even an infected person who has not yet recovered or been cured can be further infected through additional exposures (as in the case of many macroparasitic diseases, such as schistosomiasis) and if each round of superinfection does increasing additional harm.

Strategic Behavior

Together, many of the preceding considerations suggest that a distinction can be made between diseases that imply strategic behavior and those that do not. Strategic behavior arises when the contacts that put a person at risk are not anonymous and a

person can think about the information that prospective contacts have about their own infection status and their incentives to cooperate in providing such information. Issues arise of negotiation, retaliation, and incentives to tell the truth. Such interactions are conditioned by whether people are entirely self-interested or at least partially altruistic in caring about whether they infect others.

Once again, it is a question of objectives and constraints. What are the boundaries of the decisionmaking unit? Oneself? Oneself and a regular sexual partner in the case of sexually transmitted disease? Nearby neighbors in case of diseases transmitted through effluent discharge? Sexually transmitted diseases are the classic case of a strategic disease, and HIV/AIDS has put these concerns at center stage. But the practice in Japan of wearing face masks when infected by a common cold or flu virus is another example of people taking precautions to avoid infecting others, even people who are not in their immediate circle of family and friends.

Testing

There are good reasons to be tested for disease, even a noninfectious one such as cancer, but testing raises the most complex considerations when the response to a disease may be strategic, as with HIV/AIDS. Although it seems obvious that people generally want to remain uninfected and probably prefer not to infect their contacts, they may or may not want to learn their infection status, and if they do learn it they may use such knowledge for many purposes.

In the case of HIV/AIDS the decision on whether to test can be influenced by several factors: the probability that one or one's potential partners are infected; whether one cares only about oneself (egoist) or also about infecting others (altruist), something to which observed testing behavior speaks; whether one can know confidently the HIV status of one's potential partners; what one's partner might do on learning that one has been tested and perhaps, in addition, on learning one's test result, such as dissolving the partnership or taking punitive measures; the costs of being tested, including going to a test facility and coming back for test results, which may not be negligible;¹ available treatments if HIV-positive; whether legacy issues are a concern if one is infected, for instance, arrangements for one's children; whether one wants to have children and therefore unprotected sexual relations; and whether one is concerned about infecting children during pregnancy or breastfeeding.

The list is long and without clear implications. Important in thinking about tests is this question: What difference will the test result make to a person's actions afterward? Some scenarios for the case of HIV testing follow.

- People in sexual partnerships who can learn each other's infection status and who want to (and believe they can) be mutually monogamous may want to test so that they can have unprotected relations if both test negative.

- People in sexual partnerships who can learn each other's infection status may want to test so they can have protected relations or cease relations if one tests negative and the other positive, known as a discordant partnership.
- People who fear infection, are altruistic, and cannot confidently learn their partner's status may have little reason to test. If such people are HIV-negative, they take precautions to protect themselves. If HIV-positive, they take precautions to protect their partners. Such people therefore take precautions regardless of their status and, absent considerations that are not sexually strategic, such as the last four in the preceding list of motivations, have no behavior that would be conditioned on knowing their status.
- People who are egoists may increase their level of activity if they test positive because they have nothing to lose and do not care about the consequences for their partners (Philipson and Posner 1993). People viewed by moderate-risk people as too high risk to be acceptable partners who then test negative may make themselves eligible for partnerships that will sometimes infect them.

In these last cases, Philipson and Posner (1993, p. 84) argue, "testing may increase the incidence of AIDS rather than being sure to decrease it." Mechoulan (2003) shows that testing can worsen the epidemic and lower people's overall level of well-being (inclusive of their enjoyment of sexual activity) in a model of egoists who cannot find out each other's test results, a severely limiting assumption. A research priority is to understand possible outcomes when people can ask to learn the status of prospective partners and can condition their behavior on either their partner's test results or on their partner's refusal to provide results, a restriction on the potential behavior of HIV-positive egoists.

There is some statistical information on people's choices about being tested for infectious diseases involving strategic decisions, and on what they do after receiving their results. In the 1996 Tanzanian Demographic and Health Survey, 60 percent of men and 91 percent of women who reported having had a conventional sexually transmitted disease in the previous 12 months claimed to have informed their partner (Gersovitz 2002). Of the infected men, 80 percent tried to avoid infecting their partners, 15 percent did not try, and 6 percent believed their partner to be infected already. Of the infected women, 52 percent tried to avoid infecting their partners, 7 percent did not try, and 41 percent believed their partner to be infected already.² Thus, these respondents seem to be broadly altruistic in their admittedly self-reported behavior.

But the benefits of testing are very different for a conventional sexually transmitted disease than for HIV. Conventional sexually transmitted diseases are curable, and people, once cured, do not want to be reinfected by their regular partners. So rational behavior leads to possibly very different behavior in the cases of HIV and of conventional sexually transmitted disease.

Gersovitz (2002) reports some findings on HIV testing from Demographic and Health Surveys for three East African countries: Kenya 1998, Tanzania 1996, and Uganda 1995. A not insignificant number of people have been tested, from 17 percent of Kenyan men to 4 percent of Tanzanian women. In all three countries more men have been tested than women, consistent with a core group model of infection in which men infect their long-term partners after becoming infected by a small group of high-sexual-activity women (Anderson and May 1991; Over and Piot 1993). In this model of infection a negative result for a man is sufficient for the couple both to be HIV-negative. At least two-thirds of the Kenyan, Tanzanian, and Ugandan men and women who have not been tested say they would like to be, paralleling results of a survey by the Global Program on AIDS (Ingham 1995). By contrast, Fylkesnes and others (1999) report that only 1.7 percent of 4,812 randomly selected Zambian men chose to be tested and returned for their results, despite the fact that 37 percent said they were willing to be tested and counseled.

In addition to a representative random survey like the Demographic and Health Survey, there are small-scale studies of HIV testing and associated behavior based on samples of convenience that provide important hints about how people perceive their strategic situation, sometimes suggesting egoistical motivations for many. First, significant numbers of people in these studies who are tested do not return for their results, in contrast to the high percentages of respondents in the Demographic and Health Surveys who report wanting to be tested.

Furthermore, Cartoux and others (1998) report that women who tested HIV-positive were less likely to return for their test results than women who tested negative at 10 of 13 HIV testing centers in 9 poor countries.³ Perhaps these women can guess the result based on their past experiences or perhaps their current circumstances leave them less behavior to condition on test results relative to women who test negative.

Second, significant numbers of people who are tested and learn their results do not tell their partners, especially those who test HIV-positive (Ryder and others 1991; Van der Straten and others 1995; Ladner and others 1996; Grinstead and others 2001). Such behavior hints at an egoistical outlook, although the circumstances of HIV-positive women may mean that they have much less opportunity to inform their partners or much less purpose for doing so than those who are HIV-negative. For instance, they might know that their partners were their only possible source of infection.

Studies also find that people who receive HIV-positive test results do take precautions, primarily condom use or abstinence or fewer casual partners (Campbell and others 1997; Grésenguet and others 2002). Discordant partners provide an important window on whether behavior is egoistical or altruistic once infection has occurred, but the small size and potential sampling bias of some of these studies qualify any conclusions.⁴ Partners who learn that their partnership is discordant tend to adopt safer practices and to follow them more consistently than concordant partners, especially concordant negative partners (Kamenga and others 1991; Allen and others 1992; Serwadda and

others 1995; Guimarães and others 2001; Roth and others 2001). Furthermore, discordant partners tend to adopt greater precautions if the man is the HIV-negative partner (Kamenga and others 1991; Allen and others 1992; Serwadda and others 1995).

Finally, there is not much evidence on instances of violence or other punitive actions after the disclosure of test results (Maman and others 2000). Grinstead and others (2001) report that these instances were rare in their studies in Kenya, Tanzania, and Trinidad and Tobago, although Temmerman and others (1995) recount that 13 of 66 women in Nairobi suffered serious consequences following disclosures to their partners.

Aggregate Dynamics of an Epidemic, Associated Externalities, and Public Health

The aggregate dynamics of an epidemic are determined in part by biology, including the mechanisms of infection; in part by behavior at the individual level; and in part by the actions of governments. There is a lot going on, and it easily ends up a jumble. Because people as individuals do not influence the actions of governments, one way economists try to disentangle the roles of individuals and governments is to think first about the choices of individuals in the absence of governments and how these choices would aggregate. The next stage is to think about the choices that governments would make if they could hypothetically control all decisions directly in the interests of all individuals taken together, for example, by determining choices about the level of prevention and therapy for every individual without any discretion by these individuals.

If there is a discrepancy between what would happen in these two situations, that suggests the existence of externalities and a role for governments to influence the actions of individuals to transform the first situation into the second using interventions that are actually available to governments, such as financial incentives, rather than the hypothetical direct control. This section builds on the description of individual choices in the preceding section and looks first at government's constraints and objectives, then at the aggregate dynamics of infection, briefly at many different types of externalities and what governments can do to offset them, and finally in some detail at the design of immunization strategies.

Social Constraints and Objectives

Governments, like individuals, face constraints and have objectives. Much of what constrains individual behavior constrains government behavior as well, such as the biology of the disease and the associated preventive and therapeutic possibilities.

Important additional constraints on governments are the choices by individuals as conditioned by their own constraints and objectives.

There is also the issue of government objectives. Although the rationality postulate provides a basis for predicting individual behavior, it is also closely related to most economists' solution to another problem—how to evaluate what happens as an outsider or policy analyst and therefore how the government should make decisions. We started with the view that individuals choose what is best for them individually, so it is not a big stretch to assert that actions should be valued the way individuals value them.

In the context of infectious diseases the sovereignty of individual preferences means that an evaluation of an epidemic is not based solely on infection rates but takes into account that individuals also value engaging in risky activities that lead to infection. For instance, there is no point in evaluating an intervention to prevent sexually transmitted diseases solely in terms of the effect on the prevalence of the disease without recognizing that people value physical intimacy. If they did not, presumably, there would be no epidemic to begin with. The cost of a disease is the total diminution in people's well-being consequent on all the ramifications of the existence of the disease; no economist should think otherwise.

When individuals do not take into account the consequences of their actions for others, however, there are externalities. In this case individuals' actions taken together do not lead to the best overall outcome for them, opening a role for government in public health. Thus just because the government respects the sovereignty of individual preferences and therefore uses the same method to value the consequences to specific individuals as those individuals do, that does not mean that the world as it exists is the best possible world. With externalities, individuals are not fully incorporating all the social consequences of each decision when they choose among them.

The following dissection of the structure of externalities in public health shows how the characteristics of the infection determine the qualitative nature of policies. The nature and extent of externalities provide guidance on how to make a division between private and public efforts to control disease. Individuals balance the private costs and benefits as they perceive them, and governments adjust the perceived private costs and benefits to coincide with the social costs and benefits by using their power to tax and subsidize.

These considerations provide a method for analyzing such common concerns as the possibility that innovations in health technology could reduce overall well-being by indirectly promoting riskier behavior and infections, a case of immiserizing innovation. For instance, there may be a concern that new, partially effective drugs or vaccines could lower the costs to any individual of risky activity, thereby leading to more infections than without such behavior change. The discussion of testing has already raised this class of possibilities. From the perspective of the person making

the decision, the additional dangerous behavior is worthwhile given the individual benefits of the behavior and the diminished personal consequences made possible by the innovation. Concerns about immiserizing innovation, therefore, have validity only to the extent that the changed behavior has implications for people other than the person making the decision so that the externality is worsened in a way that cannot be offset by the government at sufficiently low cost.

A rationale for government intervention based on externalities does not depend on either myopia (a failure by individuals to understand or incorporate the future evolution of the epidemic into their decisions) or any other discrepancy between the way governments and individuals evaluate outcomes. Of course, if such discrepancies did exist, there could be rationales for government intervention in addition to externalities.⁵

Estimating the costs imposed on others by an infectious person is a difficult but unavoidable problem in establishing the extent of an externality. What is the monetary equivalent of the pain of illness or loss of life? Economists have some answers. One approach derives individuals' valuation of risk to health and life from the wage premium they require to do jobs that pose these kinds of risks (Viscusi 1992; Jones-Lee 1994; Pauly 1995).

Another approach avoids the explicit valuation of illness and death altogether when comparing public policies, using cost-effectiveness analysis instead to derive priorities among inputs. The costs of an input are divided by a given outcome, usually number of lives or life years saved or cases of disease prevented, yielding a result such as cost per life saved or disease prevented. Inputs with the lowest cost in these terms would then be selected. But this method can be severely misleading for a variety of reasons (Hammer 1993; Auld and others 2002). Here, the focus is on weaknesses resulting from ignoring the ways in which people affect and value health outcomes.

For one thing, cost-effectiveness can produce very unsatisfactory implicit valuations of life. Hammer (1993, p. 20) gives the following example:

Consider a situation in which two drugs are available to treat a particular disease. Drug 1 changes the probability of avoiding death from 0.2 to 0.3 and costs \$5 per treatment. Drug 2 changes the probability from 0.2 to 0.25 and costs \$2 per treatment. The cost per life saved by drug 1 is \$50 ($\$5/[0.3-0.2]$), while lives saved by drug 2 cost \$40 ($\$2/[0.25-0.20]$), making drug 2 more cost effective.

Most people probably would opt for drug 1, though, provided they are willing to pay more than \$60 to save their life. For any imputed value of life greater than \$60, the value of the increased probability of recovery outweighs the extra cost of the drug . . . Cost-effectiveness ratios, while seeming to avoid the contentious issue of deciding on a monetary value of life, merely disguise an implicit valuation that may not reflect people's preferences.

The reason cost-effectiveness leads to an unattractive choice in this example is that it implicitly and implausibly assumes that two doses of drug 2 are as effective as but cheaper than one dose of drug 1. Unfortunately, the biology of chemotherapy is unlikely to work this way.

Second, cost-effectiveness fails to consider externalities (or other market failures) in the health sector. It prescribes public spending without incorporating any criterion for distinguishing when expenditure should be in the private sector, with all the advantages of market-based incentives for efficiency, and when public involvement is needed to offset market inadequacies, such as externalities. It may be that the greatest cost-effectiveness is realized in cancer therapy. But if this therapy would be paid for by a rich private citizen on his or her own behalf, there is no obvious reason to raise tax dollars, with the associated administrative and incentive costs, to pay for such therapy through the government, and certainly no reason based on externalities.

Third, most applications of cost-effectiveness focus exclusively on benefits arising from improvements in health—diminutions in disease and death. But many health inputs may bring other benefits, such as the ability to farm land that was previously inaccessible because the risk of disease was too high.

Aggregate Dynamics

The marriage of models of rational behavior by individuals and the conventional aggregate models of epidemiology provides a basis for understanding the dynamics of the overall infection rate, the scope for externalities, and optimal public policy.

Diseases transmitted person to person. For diseases that spread from person to person, the simplest model of mathematical epidemiology determines the change in the number of new infections per period:

$$\text{New infections} = \alpha(\text{total susceptible people}) (\text{aggregate infection rate}) \quad (1)$$

The equation assumes a homogeneous population in which people contact each other randomly and with equal probability. The probability per contact of a susceptible person's meeting an infected (and infectious) person is the proportion of infected people in the population, the aggregate infection rate. In more complex situations, there may be many groups of people, each homogeneous within itself but differing in the probability with which members contact members of other groups, given that all people have one contact and each group has its own probability that a member is infected (the group's infection rate). In the homogeneous model of equation 1 the product of the aggregate infection rate and the total number of susceptible people gives the total number of susceptible people who meet an infectious person when every susceptible person has one contact. The factor α is an adjustment incorporating both the number of contacts per person per period and the inherent infectiousness

of an infected person and susceptibility of a susceptible person. Here again, in a non-homogeneous model, different groups may have different numbers of contacts per unit of time or inherent infectiousness or susceptibility.

In addition to the number of new infections a fully specified epidemiological model also determines the change in the total number of infected people as the difference between the number of new infections (equation 1) and the reduced number of infected people as they recover and become susceptible again, recover and become immune, or die. Other equations in the model determine the change in the numbers of people in other health statuses (susceptible and immune) and the change in the total population (births and deaths), and thereby determine the infection rate itself (Anderson and May 1991).

At the aggregate level the discussion of rational choice provides a justification for making the infection parameter (α in equation 1) a function of preventive measures and making the probabilities of recovery or death (in the other equations of a complete model) functions of therapeutic measures. The values of these health inputs (preventions and therapies) are the outcome of individual and government choices.

The chosen level of prevention is likely to depend on the risk of infection, providing scope for dynamic feedback in a model of choice-based epidemiology. Thus models of disease transmission that allow for endogenous behavior imply that α is a function of the infection rate (prevalence) rather than a constant as in classic epidemiology, bringing in the notion of the prevalence elasticity of prevention (Philipson and Posner 1993; Kremer 1996).⁶ A distinction between a purely epidemiological model and a choice-based one, therefore, is that in the epidemiological model the proportion of (remaining) susceptible people who become infected in each period always rises as the infection rate rises, whereas in the choice-based model it may or may not rise, depending on whether and how much people push the transmission parameter (α in equation 1) down as the infection rate rises.⁷ Philipson (2000) reviews evidence from the United States that indicates that people's preventive behavior responds to prevalence; there are no similar studies for poor countries. Choices about therapy also depend on the risk of infection and therefore on prevalence. For instance, when it is very likely that someone will become reinfected there may be little incentive to incur all the costs of a cure.

Vector transmission. Many infections are transmitted by vectors rather than directly from person to person. In these cases a person's risk of infection depends directly on the proportion of the vector that carries the infectious agent and only indirectly on the proportion of people who are infected and who may in turn infect the vectors. Consequently, even the simplest models of such diseases have more parameters than the corresponding models of diseases that are transmitted directly from person to person. These parameters represent the scope for choice by individuals and governments.

In situations of vector transmission, common parlance classifies several economically different types of inputs as prevention. One type affects the vector directly—for instance, spraying mosquitos that transmit malaria. A second type affects the probability that an uninfected person will be infected by an infected vector but does not prevent uninfected vectors from becoming infected by an infected person. These inputs break the human–vector cycle at only one point; an example is wearing waterproof footwear to avoid schistosomiasis.

A third class of inputs affects the probability that an uninfected vector will be infected by an infected person, but does not prevent an uninfected individual from becoming infected by an infected vector. An example is a program that encourages people not to void parasitic eggs into an environment where the eggs can mature. A fourth class of inputs affects both the probability that an infected person will infect an uninfected vector and the probability that an infected vector will infect an uninfected person, for example, wearing protective clothing and using bed nets to lower the chance of being bitten by a mosquito vector. In addition to these categories of prevention, there are various therapies for many vector-borne infections.

Simulations by Gersovitz and Hammer (forthcoming b) show that these different preventive and the therapeutic elements of an optimal package respond in different ways to the stage of the epidemic (aggregate rates of infection of humans and vectors and the ratio of vectors to humans) and to the underlying parameters of the model, especially the prices for the different inputs and the costs of illness. For instance, the optimal steady-state use of each input falls as its own price rises, but the response of one input to changes in the prices of other inputs varies, so that different types of inputs are not uniformly either complements or substitutes.

Furthermore, the paths of the inputs are sensitive to the choice of group targeted (the uninfected, the infected, or everyone) and to changes in how people value the sacrifice of other consumption as additional resources are transferred to the health sector. In one example, changing the targeting of a curative input from the infected to everyone reverses the direction of its path to the steady state so that use of the curative input rises rather than falls on the path to a lower steady-state level of infection from the original steady state without any inputs. Mass targeting of undiagnosed populations makes sense if it is impossible or very costly to diagnose the disease and if the costs of treatment, including side effects, are low. In the case of schistosomiasis, for example, some analysts (Prescott 1987) believe that mass undiagnosed treatment is desirable, so this type of targeting issue is a real concern.

The conclusions in Gersovitz and Hammer (forthcoming b) depend on simulations of models with parameters that have no empirical foundation and are therefore only suggestive of possibilities in the design of optimal responses to vector-borne infections. In particular, information is lacking on the relationships between the levels of prevention and therapy and their associated costs, on one hand, and the effectiveness of each level of these inputs in diminishing transmission or mitigating infections that

have occurred on the other hand. The epidemiological literature provides essentially no information on the functional forms and parameters of epidemiological processes that doubtless underlie the findings of Gersovitz and Hammer. Firmer conclusions must await a better empirical understanding of the disease processes under alternative input packages. By raising possibilities, these simulations may stimulate empirical work on these missing elements with the goal of more realistic modeling.

Externalities and the Design of Health Interventions

Externality is a central concept in formulating public health policy. If individuals are not compensated for benefits or charged for costs that they generate for others outside of family and friends, they usually do not take those consequences into account in making their decisions, leading to an externality. Economists commonly propose subsidizing actions that produce benefits for people other than the person taking the action and taxing actions that produce neglected costs. By inducing decisionmakers to take into account the effects on others, government interventions make people internalize the externality.

In principle, governments can subsidize or tax privately chosen levels of prevention and therapy to the extent that these activities affect the health of others. In practice, it may be difficult or impossible to promote certain preventive and therapeutic activities through price-based incentives. For some diseases there are inputs that are marketed, such as medicines, and inputs that involve an individuals' nonmarketed and unobservable actions, such as avoiding places and times that vectors are active or taking medicines as prescribed. Subsidy or tax interventions may be infeasible for nonmarketed and unobservable actions for which there are no transactions and therefore no prices.

For some diseases that have both marketed and nonmarketed preventive and therapeutic activities, government interventions may have to be targeted only to the marketed components. These price-based interventions will not usually achieve the best outcome. First, in some cases, even subsidizing the marketed component at 100 percent (making it free) may not imply a high enough subsidy of the total costs (marketed plus imputed nonmarketed). A greater than 100 percent subsidy is, of course, infeasible; for instance, paying people to take condoms will presumably run an infinite bill as anyone, whether a user or not, is happy to get paid to carry them away. Even when less than a 100 percent subsidy on the marketed component is in principle adequate to achieve the desired subsidy of total costs, individuals may substitute subsidized inputs for unsubsidized inputs to produce improved health, creating inefficiencies.

For other diseases that involve both marketed and nonmarketed inputs, health professionals may be able to follow up to ensure that patients use nonmarketed inputs, as in the case of directly observed treatment short-course (DOTS), for ensuring

that tuberculosis patients take their medication. This approach tries to enforce behavior by observing it, rather than relying on price-based incentives, and government spending on enforcement is justified by success in offsetting the externality.⁸ Imposing a quarantine has a similar motivation. In the case of an incurable sexually transmitted disease such as AIDS, there are very limited opportunities for government subsidies to prevent infection (on tests, treatment for curable sexually transmitted diseases that facilitate HIV infection, condoms, and clean needle programs, but not directly on safe sexual activity or clean needle use). Direct observation of behavior is essentially infeasible in most situations that spread sexually transmitted diseases, although the Thai government has implemented a campaign to enforce condom use in brothels.

As mentioned, the increasing marginal cost of successive sacrifices of nonhealth expenditures and declining marginal productivity of health inputs together often make it undesirable to take the necessary steps to eradicate a disease, even when eradication is technically possible. Although wiping out a disease is desirable in principle, people rarely find that eradication is desirable given its costs. The question becomes one of developing a policy package that is the best possible without eradicating the disease.⁹ The package involves intervening to offset externalities on an ongoing basis, accepting that the tradeoff of social costs and benefits leads to a situation in which the disease is endemic but prevalence is lower than it would be without optimal government intervention.

Diseases that involve transmission from vectors to humans and back to vectors fit the externality model well. For example, a farmer who is infected with schistosomiasis and voids one stage of the parasite into an irrigation canal where it can mature into another stage that infects other farmers probably does not incorporate these external costs fully into decisionmaking, if he even considers them at all. For this type of infectious disease the people put at risk are often neither friends nor family about whom the farmer cares most, and in any case the way the farmer provides the next link in the chain of infection is not transparent. The same situation would seem to prevail in the case of diseases transmitted directly from person to person but relatively anonymously, perhaps before infected people are even aware that they are infectious. Influenza, such as the one that killed millions of people in 1919, is a good example.

Sexually transmitted diseases seem to allow more scope for altruism because a sexual partner is rarely anonymous and is often included among people whose interests an infected person wants to take into account. Still, as the evidence suggests, many people may behave egoistically and may not take others' interests into account in important aspects of their decisions about HIV/AIDS and other sexually transmitted diseases. Finally, unlike infectious diseases of all sorts, cardiovascular disease, cancer, and some other illnesses have no externality because individuals do not put others at risk of infection, so there is no externality rationale for government intervention, although there may be other rationales, such as a failure in insurance markets (not discussed in this article).

For infectious diseases it is useful to distinguish two types of externalities, the pure infection externality and the pure prevention externality (Gersovitz and Hammer forthcoming a). The pure infection externality arises if individuals do not take into account how their becoming infected affects the risks to others of becoming infected. The pure prevention externality operates regardless of whether the individual making choices becomes infected. An example would be spraying insecticides or draining swamps, which lower others' risks of becoming infected regardless of whether the person who pays for the preventive actions becomes infected.

This distinction between externalities provides some qualitative guidance for government interventions. Consider a quintessential infection externality: The disease is transmitted from person to person; people are either infected and infectious or well but susceptible to infection; and once cured, a person is again susceptible. In offsetting the externality, it is equally desirable to prevent someone from becoming infected as to cure them: The goal is to keep more people out of the infectious group than would otherwise be the case. Consequently, prevention and cures should be subsidized at exactly the same rate, a good instance when the two types of inputs are identical from an economic viewpoint (Gersovitz and Hammer forthcoming a). This result is not general but depends on the structure of the disease as defined by what happens to an infected individual. Recovery to susceptibility, recovery to immunity, or death all imply a different distribution of government intervention between prevention and therapy.

Vector-transmitted diseases provide further complications in the design of an intervention package.¹⁰ Interventions are not restricted to activities that either prevent people from infecting others or that treat the infected. Some inputs affect the vector and its ability to infect people. In the case of vector control, pure prevention externalities arise to an extreme degree. The government may have to pay for the whole program of vector control because control at the individual level may be impractical or ineffective; the infiltration of vectors from outside a single individual's perimeter of control may be overwhelming. For instance, spraying in one's own compound may have little effect on the probability of getting malaria if one's neighbors do nothing. Without a method of coordinating large groups of people, the definition of a government, the result will be that no one bothers to spray.

Miguel and Kremer (2001) estimate the externalities associated with a curative input—deworming school children in Kenya. The positive externalities to some children of treating other infected children may be large enough to justify paying children to be dewormed over and above paying all the costs of treating them. Miguel and Kremer also emphasize the need to incorporate the consequences of externalities in estimating the benefits of treatment if the treatment group and the associated external benefits are large enough to improve the health of untreated individuals in the control group who live in the same community. In this case the simple difference between the outcomes of randomly selected treated and untreated individuals would

understate the benefits of treatment to treated individuals and would entirely neglect the indirect benefits to the untreated.

Although the general notion of externalities is straightforward, seemingly paradoxical results are possible, especially when people differ in number of contacts and probabilities of being infectious (nonhomogeneous groups). For instance, consider a disease that spreads in a population made up of two groups, one with a high exposure to infection and one with a low exposure. For concreteness, say that the infection is HIV/AIDS and that the first group consists of people with many sexual partners and the second of people with relatively few partners. From an individual's perspective there is no doubt that lowering the number of partners will lower the probability of becoming infected, other things equal. But what happens to the infection rate in the population as a whole?

Assume initially that people in the low sexual activity group (as measured by number of partners) are not sexually active. In this case the epidemic may spread rapidly among people in the high-sexual-activity group. Now consider an increase in the activity level of people in the low-activity group, perhaps to one contact. Some of these contacts will be with people in the high-activity group and some will result in infections of people in the low-activity group. But if the people in the low-activity group have only one contact, they will not infect anyone else. In the meantime, some of the people in the high-activity group who are not yet infected will be less likely to meet infected people in the high-activity group because some of their contacts will be with people in the low-activity group, so they will be less likely to be infected. In principle, this effect could be so strong as to extinguish the epidemic among people in the high-activity group and therefore among the population as a whole, although at a cost to people in the low-activity group who become infected during the transition.¹¹

This example illustrates a particularly strong positive externality of increased activity by people in the low-activity group. Or, to turn this lesson around, the observation of an increased prevalence among people in the high-activity group need not reflect an increase in risky behavior by people in this group. Instead, it may reflect decreased risky behavior by people in the low-activity group. Indeed, high-activity people may even be decreasing their risky behavior somewhat.

Spurred by this type of stark and simplified example, Kremer and Morcum (1998) looked for this effect in a more realistic model of HIV transmission calibrated to data from the United Kingdom. They estimate that 8 of 10 heterosexual people have sufficiently low sexual activity levels (as measured by number of partners) that an increase in their sexual activity would decrease the national prevalence level. Thus, the theoretical result may not simply be a curiosity.

Understanding this phenomenon may be important for interpreting the dynamics of an epidemic with different groups (Whitaker and Rentin 1992), but its policy implications are less clear. Few people would want to encourage people with low

sexual activity to take risks with their health for some notion of the social good. Nonetheless, such a finding would seem to suggest the need to find policies to lower infection rates among people in high-sexual-activity groups who may actually be made more vulnerable by campaigns that decrease activity among people in low-activity groups.

Immunization: A Surprisingly Controversial and Important Source of Externalities

Public immunization programs are often cited as an example of policies to overcome externalities, precisely because people do not take into account the benefits arising from their acquired inability to infect others. Stiglitz (2000, p. 134) presents the arguments for this position: "While the main beneficiary of a vaccination may be the individual protected, and there is a significant marginal cost of vaccinating an additional individual, the public health benefits from universal vaccination—the reduced incidence of the disease, possibly its eradication—are benefits from which no one can be excluded."

Recently, Francis (1997) and Philipson (2000) have disputed this conventional view. Thus Philipson (2000, pp. 1763–64) remarks that:

relying on standard arguments about the positive external effects of disease prevention, economists often [argue] for an active public role in the prevention of infectious diseases, such as AIDS (see for instance Stiglitz 1997, p. 15). However, economists have rarely attempted to explain patterns of disease occurrence or to evaluate public interventions in the context of a society with individuals who do the best they can given their constraints. Such recent analysis...has cast...doubt on the old textbook arguments by economists.

The recent analysis on which Philipson makes his case, however, does not really support the abandonment of Stiglitz's and others' intuition because it is based on very narrow assumptions.

Francis (1997) provides a formal example of a dynamic model of infection propagation in which there is no externality, or at least none that justifies government intervention. Francis's assumptions are deliberately chosen to provoke further research. He assumes that there are no births of uninfected nonimmune people, that no one recovers from the disease, and that no one dies of any cause.

Gersovitz (2003) gives the following explanation for Francis's conclusions. When there are no births or deaths and no one is vaccinated, infection spreads (as determined by equation 1), the proportion of susceptible individuals falls continuously and the proportion of infected individuals rises continuously. As the proportion of infected individuals rises, so do the risk of infection and the benefits of immunity through vaccination for the remaining susceptible individuals. Eventually the

proportion of the population that is infected (and therefore the probability of infection) reaches a threshold at which the value of being vaccinated just equals the cost of vaccination (which is constant). If it is worth vaccinating anyone it is now worth vaccinating every susceptible individual because the proportion of infected individuals (and the probability of infecting the remaining susceptible individuals) will otherwise continue to rise; it cannot fall because no new susceptible individuals are born and no infected individuals die. The only way to maintain the steady-state value of the infection rate is to vaccinate every susceptible individual once the threshold is reached. In the steady state the proportion of susceptible individuals is therefore zero.

The form of the rule for vaccination is therefore a threshold rule and is the same whether governments or private individuals decide on vaccination. Furthermore, the value of the threshold is the same, and in this sense there is no externality, at least not one that justifies intervention. It might be tempting to infer that government would want to set a lower threshold that is attained earlier. But why? Without government intervention, once the private threshold is reached, everyone gets vaccinated anyway, and if everyone is vaccinated when anyone is, there is no one left to benefit from earlier vaccination, which would keep down the aggregate infection rate, so there is no point to the government's inducing an earlier time for vaccinating everyone. Hence the values of the thresholds are the same.

Gersovitz (2003) also considers the case in which there are new births and people die, but only as a proportion of the people who are alive so that any group of people would die out asymptotically, a conventional assumption made for tractability in mathematical epidemiology. The assumption of births and deaths is the only change from Francis's model; infected individuals are still assumed never to recover. In this case there are people, the newly born, who would benefit from a lower threshold infection rate for vaccination that is reached earlier so that when they enter the world the infection rate is as low as possible. Their elders decide about vaccination based only on their own benefits and costs, ignoring the benefit to newborns, even their own, because they take the aggregate infection rate as external to their own decisions about vaccination. After all, each member of the elder cohort is only one of many people who influence the aggregate infection rate. The government, however, recognizes that if the threshold were lowered, newborns will benefit and that it can induce a lower threshold by subsidizing immunization. In addition, once the government's threshold infection rate is reached, not everyone who is susceptible is vaccinated.

An assertion that it is optimal to vaccinate everyone when the threshold is reached would involve a contradiction: If a rule of universal immunization were adopted, all new susceptible individuals would be vaccinated, thereby diluting the infected pool as a proportion of the total population and lowering the infection rate below the threshold. But this conclusion contradicts the notion of a threshold infection rate after which everyone is vaccinated. Therefore the government's optimal

strategy is to begin vaccinating a proportion of susceptible individuals less than one once a threshold infection rate has been reached and to keep vaccinating members of the susceptible pool inclusive of newborns in such a way that the proportions of susceptible, vaccinated, and infected people remain constant.

Without government subsidization of immunization to equalize the two thresholds, private individuals follow the same qualitative rule of a threshold, but the threshold is higher because it does not incorporate the impact on the aggregate probability of others becoming infected. In effect, births introduce heterogeneity into the population: Not everyone is alive at the same time, and newborns cannot take decisions simultaneously with those born before them. The result of a positive subsidy obtains whether population growth is positive or not; all that is needed is that the birth rate be positive, not that it exceed the death rate.

Gersovitz (2003) also shows that if infected individuals recover and become susceptible again, recovery can play the same role as a positive birth rate in generating an externality and a motive for government intervention. For both births and recoveries there is a straightforward formula for the optimal subsidy of the cost of vaccination to internalize the externality. The absolute value of the subsidy depends negatively on the cost of the vaccination, the interest rate, and the natural death rate of the population and positively on the cost of infection, the birth rate of the population, and the recovery rate, but not at all on α , the transmission parameter of equation 1.

The purpose of government intervention is to improve welfare by lowering endemic prevalence. Only if the cost of infection is infinite or the cost of immunization is zero does the government's optimal threshold for beginning immunization equal zero so that the disease is eradicated (but only asymptotically because the population of infected individuals only dies out asymptotically, a modeling artifact of the assumption made for mathematical tractability that a constant fraction of the population dies in every period). In this case of an infinite cost of the disease or a zero cost of immunization, the government threshold coincides with the private threshold, and because both thresholds are set to zero there is no need for government intervention to bring about eradication.

Contrast these results on subsidization and eradication with those of Geoffard and Philipson (1997, p. 222), who believe that the inability to achieve eradication "brings into doubt classic justifications of Pigouvian subsidies aimed at solving the underprovision of vaccines due to their positive external effects."¹² A policy of eradication is not synonymous with a policy that optimally offsets the externality by using subsidies. Although eradication may well be too expensive relative to its benefits, there is still a rationale for the optimal subsidization of immunization to lessen the endemic prevalence of disease. Indeed, with current models dependent on the assumption that disease can only die out asymptotically, a research priority is modeling immunization when eradication is feasible in finite time.¹³

At this stage, it seems that the Francis-Philipson position against an externality-based rationale for vaccination does not extend beyond one generally inapplicable example, really more of a curiosity and a stimulus to further research than a guide to policy. More work, however, is needed to delineate the full scope of externalities associated with vaccinations.

Conclusion

This article has examined the implications of rational choice for both individual behavior and public policy. It argues that this way of looking at how people make decisions that affect their health helps in understanding what people do and what governments can and should do to change health outcomes. The structure of objectives and constraints is critical in building this framework. Biology and epidemiology are only two components of this framework but are important ones for determining constraints.

At the individual level models of rational choice generate hypotheses about what people do about their health, especially how they respond to increased risk of infection. This approach provides insight into rational fatalism, cases in which individuals react to increases in the probability of infection by decreasing certain types of preventive behavior. It also yields a typology of the consequences of the availability of tests based on the circumstances of individuals.

The greatest relevance of these models for public policy is for explicating the aggregate consequences of people's individual choices. Here the concept of externalities comes to the fore. Models of infectious diseases are rife with externalities, and the conclusion is that choices by individuals will almost never lead to the best outcome based on social criteria, even when the social criteria respect the sovereignty of individual preferences.

But the literature on externalities and infectious diseases does not stop with the general statement that private choices do not produce socially optimal outcomes. It provides guidance on how the government should intervene, usually through subsidies, to improve people's well-being as measured from a social perspective. It provides qualitative results on how interventions should affect incentives for prevention relative to therapy and, when there are many different types of prevention, on how to adjust the incentives to undertake them. It provides results on when and how to subsidize vaccinations as a function of such underlying parameters as the costs of illness and vaccination, the birth rate of susceptible individuals, and the recovery rate of infected individuals. These and other conclusions are a consequence of marrying economics and epidemiology.

For this marriage to be as productive as possible, however, more will be needed than simply a combination of what each side has already brought to the union.

Full consummation of the union requires quantitative information on the response of health outcomes to various preventive and therapeutic inputs. It also requires information on the choices individuals make when faced with different incentives. Efforts by both parties to establish this information should prove very fruitful.

Notes

Mark Gersovitz is Professor of Economics at the Johns Hopkins University. Jeffrey S. Hammer is Lead Economist, Development Research Group at the World Bank. Parts of this article draw on Gersovitz's (1999) Joseph Fisher Lecture at the University of Adelaide, in turn the basis for Gersovitz (2000).

1. For instance, when an HIV/AIDS testing and counseling center in Kampala, Uganda, replaced tests that required clients to return in two weeks with same-day results, the share of people tested who learned their test results rose from 79 percent to 100 percent (Kassler and others 1998). On the other hand, it may be that the additional 21 percent who learned their results had little or no use for the knowledge and that costs were not a deterrent.

2. Petry and Kingu (1996, p. 267–268) report that almost all people in their sample with a conventional sexually transmitted disease were aware of it, although only 25 percent of patients who reported sexually transmitted disease–like symptoms had a conventional sexually transmitted disease.

3. Among more recent studies with relatively large samples, Kawichai and others (2002) report that HIV-positive men were less likely and HIV-positive women equally likely to return for their tests as the corresponding HIV-negative groups. Grésenguet and others (2002) report no difference in serostatus between people who do and do not return for test results.

4. The ratio of male HIV-positive to female HIV-positive discordant partnerships is close to one in studies that report the serostatus of discordant couples (Gersovitz 1999). These results include: 69 male HIV-positive and 80 female HIV-positive in Kamenga and others (1991); 12 and 16 in Van der Straten and others (1995); 44 and 35 in Serwadda and others (1995); approximately 350 of each in McKenna and others (1997); 18 and 39 in Grinstead and others (2001); 43 and 23 in Roth and others (2001) and 22 and 21 in Hugonnet and others (2002). Carpenter and others (1999) present evidence that the probability of seroconversion of women in discordant partnerships is twice that of men. Even taking account of this finding, these ratios of male HIV-positive to female HIV-positive discordant partnerships suggest that a core-group model in which men become infected by very high-activity women and then infect their regular partners cannot be the whole story; too many women in the discordant regular partnerships are the infected partner.

5. Clear isolation of the dynamics of disease control and of externalities requires careful attention to the specification of private behavior. Thus Kremer (1996) posits that people care only about the chance of ever being infected and not about when they are infected. Such a utility function may be a useful first approximation, but it has clear limitations in conflicting with the criterion of present-discounted valuation of costs and benefits and therefore provides a poor basis for trading off expenditures for the control of epidemics against their (often future) benefits. Similarly, Geoffard and Philipson (1997) specify a demand function for vaccines, but only on the assumption that people choose to be vaccinated at birth or not at all and there is no discounting.

6. Most economists would find the notion of prevalence-elastic choices about prevention to be second nature. In a model of HIV/AIDS in a closed population with neither births nor deaths, however, Geoffard and Philipson (1996, p. 604) push the consequences of prevalence elasticity further than seems justified: "A growing prevalence and the public subsidy *compete* to induce protective activity, and this makes the timing of the public intervention a crucial factor in determining its economic efficiency. If the subsidy is not fast enough, the prevalence has made it irrelevant in inducing protection" (emphasis in original). This statement seems to mean that if government intervention has been delayed

sufficiently, there is no point in doing anything. But as argued later, however, externalities justify a government subsidy to induce prevention and achieve a lower rate of infection than would private behavior alone. Even if the government has waited and the infection rate has gone beyond what it should have been limited to, this situation should be only temporary. With newly born (or newly sexually active) people entering the pool of susceptible people all the time, the government can induce additional prevention and get the infection rate back down to the socially optimal level. As in many situations in life, better late than never.

7. In the case of the HIV/AIDS epidemic in Africa, traditional epidemiological models (for example, United Nations and World Health Organization 1991) have tended to underpredict the epidemic, presenting a major puzzle for understanding the dynamics of this epidemic. The possibility of a high prevalence elasticity—stressed by Philipson and Posner (1993) but omitted from traditional epidemiological models—and the consequent fall in α and the proportion of susceptible people who become infected as the epidemic evolves is not the solution. Indeed to the extent that highly prevalence-elastic behavior is operative, the puzzle is all the larger. Even if rational fatalism were operative, the actual α would be higher relative to the epidemiologists' (constant) estimate only if α were estimated from situations when infection rates were already fairly high but before rational fatalism became operative. If instead α were estimated from the incipient stages of the epidemic, rational fatalism could never lead to an actual α higher than the initial estimate because even rationally fatalistic people would not want to engage in more sexual activity than when there was almost no risk. In models of rational fatalism the only benefit people obtain from sexual activity is the activity itself, which is traded off against the (undesirable) risk of becoming infected. Within the confines of these models, the only people who would want to have more partners in the presence than in the absence of the disease would be people who wanted to engage in more risky behavior precisely because they valued the chance of becoming infected. A different set of considerations arises if one steps outside these models and thinks about people for whom sexual activity is part of relationships and family formation that imply fewer partners than people would otherwise prefer. If the epidemic were somehow to disrupt family formation with its associated restraint on the number of partners, then it is possible that sexual activity would increase as the epidemic progresses, but such speculation is outside the currently available models.

8. A person who gets cured provides a benefit to others whom he or she might have infected, but a person who gets cured in such a way as to induce a disease's resistance to drugs has a negative effect on others who become infected by making it harder for them to be cured. Traditional medical practice is to keep prescribing the best among a group of antibiotics until resistance to it arises and then to move to the next best and so on, "best" being determined by some combination of cost and efficacy. A number of economists have argued, however, that more than one antibiotic should be used simultaneously on the patient population (although not on any individual patient) and that doing so will minimize the costs of resistance (Laxminarayan and Brown 2001; Laxminarayan and Weitzman 2002; Laxminarayan 2003). By contrast, Salant (2003) emphasizes the difficulty of treating similar patients differently in the name of the social good and thereby questions the practicality of a multidrug strategy of minimizing drug resistance. Nonetheless, the negative externality associated with resistance would seem to justify unambiguously a public subsidy of programs, such as DOTS to ensure that people comply with drug regimens even under the traditional one-drug-at-a-time strategy, because noncompliance is a source of drug resistance.

9. Much of the literature that deals with the optimal use of inputs in mathematical epidemiology formulates the problem in such a way that inputs are used to the fullest extent or not at all. The reason is that this literature ignores the decreasing marginal product of inputs and the increasing marginal utility of the diverted resources that are necessary to pay for the inputs. Wickwire (1977) surveys early work; Sethi and Staats (1978), Greenhalgh (1988), and Hocking (1991) are more recent. The economists Geoffard and Philipson (1996, 1997) are similarly concerned primarily with the question of whether governments can eradicate a disease and not with whether it is optimal to do so or to choose instead among endemic rates of infection. The only mathematical discussion of when eradication dominates chronic control of an infection (and vice versa) that we know is by the economists Goldman and Lightwood (2002). Their analysis is of a somewhat special case because it only looks at

therapies, ignores the increasing marginal utility of diverted resources and uses a model in which the disease can only die out asymptotically. Nonetheless, it should be of pivotal importance to understanding this subject.

10. In the case of schistosomiasis, Wiemer (1987) looks at more than one input at a time, but the specification of costs and benefits implies that one of two inputs is used either to the maximum extent possible or not at all, a rather special case, and there is no discussion of externalities and the division of decisions between the private and public sectors.

11. Kremer (1996) shows that there can be multiple equilibria when people choose their levels of sexual activity endogenously and differ in their preferences about activity and the risk of infection.

12. Note that in the model of Geoffard and Philipson (1997) death is proportional to the population so that any group in the population never dies out in finite time. Consequently, eradication in their model is at best asymptotic. Furthermore, they do not provide a discussion of the social optimality of eradication under their assumption of heterogeneous demand for vaccines, so it is hard to contrast their findings directly with Gersovitz (2003). Finally, in their brief discussion of eradication, when the disease presumably may die out in finite time, Geoffard and Philipson (1997, p. 227) do not analyze when eradication is optimal but only provide weak upper and lower bounds on the costs and benefits of eradication. (See also the discussion in note 6.)

13. An overlapping generations (OLG) model is one way to develop this strategy; Mechoulan (2003) uses this approach in his analysis of testing.

References

- Allen, S., J. Tice, P. Van de Perre, A. Serufilira, E. Hudes, F. Nsengumuremyi, J. Bogaerts, C. Lindan, and S. Hulley. 1992. "Effect of Serotesting with Counselling on Condom Use and Seroconversion among HIV Discordant Couples in Africa." *British Medical Journal* 304:1605–9.
- Anderson, R. M., and R. M. May. 1991. *Infectious Diseases of Humans*. Oxford: Oxford University Press.
- Auld, C. M., C. Donaldson, C. Mitton, and P. Shackley. 2002. "Health Economics and Public Health." In R. Detels, J. McEwen, R. Beaglehole, and H. T. Tamaka, eds., *Oxford Textbook of Public Health*, 4th ed. Oxford: Oxford University Press.
- Campbell, C. H., M. E. Marum, M. G. Alwano-Edyegu, B. Dillon, M. Moore, and E. Gumisiriza. 1997. "The Role of HIV Counseling and Testing in the Developing World." *AIDS Education and Prevention* 9:92–104.
- Carpenter, L. M., A. Kamali, A. Ruberantwari, S. S. Malamba, and J. A. Whitworth. 1999. "Rates of HIV-1 Transmission within Marriage in Rural Uganda in Relation to the HIV Sero-Status of the Partners." *AIDS* 13:1083–89.
- Cartoux, M., N. Meda, P. Van de Perre, M. L. Newell, I. de Vincenzi, and F. Dabis. 1998. "Acceptability of Voluntary HIV Testing by Pregnant Women in Developing Countries: An International Survey." *AIDS* 12:2489–93.
- Francis, P. J. 1997. "Dynamic Epidemiology and the Market for Vaccinations." *Journal of Public Economics* 63:383–406.
- Fylkesnes, K., A. Haworth, C. Rosenvard, and P. Kwapa. 1999. "HIV Counselling and Testing: Over-emphasizing High Acceptance Rates a Threat to Confidentiality and the Right Not to Know." *AIDS* 13:2469–74.
- Geoffard, P.-Y., and T. Philipson. 1996. "Rational Epidemics and Their Public Control." *International Economic Review* 37:603–24.
- . 1997. "Disease Eradication: Private versus Public Vaccination." *American Economic Review* 87:221–30.

- Gersovitz, M. 1999. *Human Behaviour and the Transmission of Infectious Disease: An Economist's Perspective*. The Forty-Seventh Joseph Fisher Lecture in Commerce. Adelaide: University of Adelaide.
- . 2000. "A Preface to the Economic Analysis of Disease Transmission." *Australian Economic Papers* 39:68–83.
- . 2002. "The African HIV Epidemic as Seen through the Demographic and Health Surveys." Mimeo, The Johns Hopkins University, Department of Economics, Baltimore, Md.
- . 2003. "Births, Recoveries, Vaccinations and Externalities." In R. J. Arnott, B. Greenwald, R. Kanbur, and B. Nalebuff, eds., *Economics for an Imperfect World: Essays in Honor of Joseph Stiglitz*. Cambridge, Mass.: MIT Press.
- Gersovitz, M., and J. S. Hammer. Forthcoming a. "The Economical Control of Infectious Diseases." *Economic Journal*.
- . Forthcoming b. "Tax/Subsidy Policies toward Vector-Borne Infectious Diseases." *Journal of Public Economics*.
- Goldman, S. M., and J. Lightwood. 2002. "Cost Optimization in the SIS Model of Infectious Disease with Treatment." *Topics in Economic Analysis & Policy* 2(1).
- Greenhalgh, D. 1988. "Some Results on Optimal Control Applied to Epidemics." *Mathematical Biosciences* 88:125–58.
- Grésenguet, G., J. Sehonou, B. Bassirou, J. D. Longo, J. E. Malkin, T. Brogan, and J. Belec. 2002. "Voluntary HIV Counseling and Testing: Experience among the Sexually Active Population in Bangui, Central African Republic." *Journal of Acquired Immune Deficiency Syndromes* 31:106–14.
- Grinstead, O. A., S. E. Gregorich, K. H. Choi, and T. Coates. 2001. "Positive and Negative Life Events after Counselling and Testing: The Voluntary HIV-1 Counselling and Testing Efficacy Study." *AIDS* 15:1045–52.
- Guimarães, M. D. C., C. Boschi-Pinto, and E. A. Castilho. 2001. "Safe Sexual Behaviour among Female Partners of HIV-Infected Men in Rio de Janeiro, Brazil." *International Journal of STD & AIDS* 12:334–41.
- Hammer, J. S. 1993. "The Economics of Malaria Control." *World Bank Research Observer* 8:1–22.
- Hocking, L. M. 1991. *Optimal Control: An Introduction to the Theory with Applications*. Oxford: Clarendon Press.
- Hugonnet, S., F. Moshia, J. Todd, K. Mugeye, A. Klokke, L. Ndeki, D. Ross, H. Grosskurth, and R. Hayes. 2002. "Incidence of HIV Infection in Stable Sexual Partnerships: A Retrospective Cohort Study of 1802 Couples in Mwanza Region, Tanzania." *Journal of Acquired Immune Deficiencies Syndromes* 30:73–80.
- Ingham, R. 1995. "AIDS: Knowledge, Awareness, and Attitudes." In J. Leland and B. Ferry, eds., *Sexual Behaviour and AIDS in the Developing World*. London: Taylor and Francis.
- Jones-Lee, M. W. 1994. "Safety and the Saving of Life: The Economics of Safety and Physical Risk." In R. Layard and S. Glaister, eds., *Cost-Benefit Analysis*. Cambridge: Cambridge University Press.
- Kamenga, M., R. Ryder, M. Jingu, and others. 1991. "Evidence of Marked Sexual Behavior Change Associated with Low HIV-1 Seroconversion in 149 Married Couples with Discordant HIV- Serostatus: Experience at an HIV Counselling Center in Zaire." *AIDS* 5:61–67.
- Kassler, W. J., M. G. Alwano-Edyegu, E. Marum, B. Biryahwaho, P. Kataaha, and B. Dillon. 1998. "Rapid HIV Testing with Same-Day Results: A Field Trial in Uganda." *International Journal of STD & AIDS* 9:134–38.
- Kawichai, S., D. D. Celentano, R. Chaifongsri, K. E. Nelson, K. Srithanaviboonchai, C. Natpratan, and others. 2002. "Profiles of HIV Voluntary Counseling and Testing of Clients at a District Hospital, Chiang Mai Province, Northern Thailand, from 1995 to 1999." *Journal of Acquired Immune Deficiencies Syndromes* 30:493–502.
- Kremer, M., 1996. "Integrating Behavioral Choice into Epidemiological Models of AIDS." *Quarterly Journal of Economics* 111:549–73.

- Kremer, M., and C. Morcum. 1998. "The Effect of Changing Sexual Activity on HIV Prevalence." *Mathematical Biosciences* 151:99–122.
- Ladner, J., V. Leroy, P. Msellati, M. Nyiraziraje, A. De Clercq, P. Van de Perre, and F. Dabis. 1996. "A Cohort Study of Factors Associated with Failure to Return for HIV Post-Test Counselling in Pregnant Women: Kigali, Rwanda, 1992–1993." *AIDS* 10:69–75.
- Laxminarayan, R., ed., 2003. *Battling Resistance to Antibiotics and Pesticides*. Washington, D.C.: Resources for the Future.
- Laxminarayan, R., and G. M. Brown. 2001. "Economics of Antibiotic Resistance: A Theory of Optimal Use." *Journal of Environmental Economics and Management* 42:183–206.
- Laxminarayan, R., and M. L. Weitzman. 2002. "On the Implications of Endogenous Resistance to Medications." *Journal of Health Economics* 21:709–18.
- Maman, S., J. Campbell, M. D. Sweat, and A. C. Gielen. 2000. "The Interactions of HIV and Violence: Directions for Future Research and Interventions." *Social Science and Medicine* 50:459–78.
- McKenna, S. L., G. K. Muyinda, D. Roth, M. Mwali, N. N'gandu, A. Myrick, F. H. Priddy, V. M. Hall, A. A. von Lieven, J. R. Sabatino, and S. Allen. 1997. "Rapid HIV Testing and Counseling for Voluntary Testing Centers in Africa." *AIDS* 11:S103–10.
- Mechoulan, S. 2003. "HIV Testing: A Trojan Horse." Mimeo, University of Toronto, Department of Economics.
- Miguel, E., and M. Kremer. 2001. *Worms: Education and Health Externalities in Kenya*. NBER Working Paper 8481, National Bureau of Economic Research, Cambridge, Mass.
- Over, M., and P. Piot. 1993. "HIV Infection and Sexually Transmitted Diseases." In D. T. Jamison, W. H. Mosley, A. R. Measham, and J. L. Bobadilla, eds. *Disease Control Priorities in Developing Countries*. New York: Oxford University Press.
- Pauly, M. V. 1995. "Valuing Health Care Benefits in Money Terms." In F.A. Sloan, ed., *Valuing Health Care*. Cambridge: Cambridge University Press.
- Petry, K. U., and H. Kingu. 1996. "HIV Infection among Pregnant Women in Linda, Tanzania, 1989–1993." *International Journal of STD & AIDS* 7:265–68.
- Philipson, T. J. 2000. "Economic Epidemiology and Infectious Diseases." In A. J. Cuyler and J. P. Newhouse, eds., *Handbook of Health Economics*. Amsterdam: North Holland.
- Philipson, T. J., and R. A. Posner. 1993. *Private Choices and Public Health: The AIDS Epidemic in an Economic Perspective*. Cambridge, Mass.: Harvard University Press.
- Prescott, N. M. 1987. "The Economics of Schistosomiasis Chemotherapy." *Parasitology Today* 3:21–24.
- Roth, D. L., K. E. Stewart, O. J. Clay, A. van der Straten, E. Karita, and S. Allen. 2001. "Sexual Practices of HIV Discordant and Concordant Couples in Rwanda: Effects of a Testing and Counselling Programme for Men." *International Journal of STD & AIDS* 12:181–88.
- Ryder, R. W., V. L. Batter, M. Nsuami, N. Badi, L. Mundele, B. Matela, M. Utshudi, and W. L. Heyward. 1991. "Fertility Rates in 238 HIV-1-Seropositive Women in Zaire Followed for 3 Years Post-Partum." *AIDS* 5:1521–27.
- Salant, S. W. 2003. "Same Infection, Same Time, Same Antibiotic?" In R. Laxminarayan, ed., *Battling Resistance to Antibiotics and Pesticides*. Washington, D.C.: Resources for the Future.
- Serwadda, D., R. H. Gray, M. J. Wawer, R. Y. Stallings, N. K. Sewankambo, J. K. Konde-Lule, B. Lainjo, and R. Kelly. 1995. "The Social Dynamics of HIV Transmission as Reflected through Discordant Couples in Rural Uganda." *AIDS* 9:745–50.
- Sethi, S. P., and P. W. Staats. 1978. "Optimal Control of Some Simple Deterministic Epidemic Models." *Journal of the Operational Research Society* 29:129–36.
- Stiglitz, J. E. 1997. "Introduction." In World Bank, *Confronting AIDS*. New York: Oxford University Press.

- . 2000. *Economics of the Public Sector*. New York: Norton.
- Temmerman, M., J. Ndinya-Achola, J. Ambani, and P. Piot. 1995. "The Right Not to Know HIV-Test Results." *Lancet* 395:969–70.
- Van der Straten, A., R. King, O. Grinstead, A. Serufilira, and S. Allen. 1995. "Couple Communication, Sexual Coercion and HIV Risk Reduction in Kigali, Rwanda." *AIDS* 9:935–44.
- Viscusi, W. K. 1992. *Fatal Tradeoffs: Public and Private Responsibilities for Risk*. New York: Oxford University Press.
- United Nations and World Health Organization. 1991. *The AIDS Epidemic and its Demographic Consequences*. New York and Geneva: United Nations and World Health Organization.
- Whitaker, L., and A. M. Rentin. 1992. "A Theoretical Problem of Interpreting the Recently Reported Increase of Homosexual Gonorrhea." *European Journal of Epidemiology* 8:187–91.
- Wickwire, K. 1977. "Mathematical Models for the Control of Pests and Infectious Diseases: A Survey." *Theoretical Population Biology* 11:182–238.
- Wiemer, C. 1987. "Optimal Disease Control through Combined Use of Preventive and Curative Measures." *Journal of Development Economics* 25:301–19.

