Optimal Subsidies for Prevention of Infectious Disease

Matthew Goodkin-Gold
Harvard University. Email: mgoodkingold@g.harvard.edu

Michael Kremer
Harvard University, National Bureau of Economic Research (NBER). Address: Department of Economics, Harvard University, Littauer Center 207, Cambridge MA 02138. Email: mkremer@fas.harvard.edu.

Christopher M. Snyder
Dartmouth College and NBER. Address: Department of Economics, Dartmouth College, 301 Rockefeller Hall, Hanover NH 03755. Email: chris.snyder@dartmouth.edu.

Heidi Williams
Stanford University and NBER. Address: 579 Jane Stanford Way, Office 323, Stanford CA 94305. Email: hwill@stanford.edu.

Abstract: Most economists would agree that the positive externalities caused by prevention of infectious disease create a *prima facie* case for subsidies. However, little is known about the appropriate magnitude of these subsidies, or about whether the level of such subsidies should vary across diseases. We integrate a standard epidemiological model with an economic model of consumer and producer behavior to address these questions. Across a continuum of market structures, we find that the equilibrium steady-state marginal externality is non-monotonic in disease transmissiblity, peaking when the disease is just transmissible enough to survive in steady-state. This pattern implies that marginal externalities—and, as we show, optimal subsidies—are higher for serious but rare diseases relative to diseases with lower individual burden but higher disease prevalence. Crude calibrations suggest that optimal subsidies for technologies such as vaccines, condoms, and mosquito nets, which prevent infectious diseases, may be very large relative to current levels.

Journal of Economic Literature codes: O31, L11, I18, D42

Keywords: vaccine, epidemiology, externality, pharmaceutical

Acknowledgements: We are grateful for helpful comments from seminar participants at the Infectious Diseases in Poor Countries and the Social Sciences Conference at Cornell, the DIMACS Game Theoretic Approaches to Epidemiology and Ecology Workshop at Rutgers, the Federal Trade Commission's Bureau of Economics Roundtable on the Economics of the Pharmaceutical Industry, Harvard, the U.S. National Institutes of Health Models of Infectious Disease Agent Study Group at the Hutchinson Cancer Research Center in Seattle, the American Economic Association Economics of Infectious Disease session, and the Yale School of Medicine. Maya Durvasula and Frank Schilbach provided helpful research assistance. Williams gratefully acknowledges financial support from NIA grant number T32-AG000186 to the NBER.

1. Introduction

Technologies to prevent infectious diseases—such as vaccines, condoms, and mosquito nets—can generate positive health externalities. While standard economic models provide a justification for public subsidies of such preventive technologies, existing models provide little guidance on the appropriate magnitude of these subsidies. This gap in understanding makes it difficult for economists to provide guidance—even at a conceptual level—on both the optimal level of government subsidies for infectious disease control and the ways in which the level of such subsidies should vary across diseases. The COVID-19 pandemic has brought to light the profound importance of infectious disease policy, and, as can be seen in the explosion of new research, the potential for the tools of economics to inform this response.[1]

In this paper, we construct a tractable integrated economic and epidemiological model that allows us to shed light on these questions. We focus the model on a vaccine market, although the analysis is applicable to any market for a technology that prevents consumers from contracting an infectious disease. For a continuum of market structures—ranging from perfect competition to monopoly, and a more general Cournot case that nests both extremes—the model generates closed-form solutions for the equilibrium steady-state marginal externality from additional immunizations and the minimal subsidy required to achieve the first best level of vaccination. The marginal externality and minimal subsidy are functions of a key set of estimable parameters such as disease transmissibility, vaccine efficacy, and cost; we report some—admittedly crude—calibrations for particular diseases.

A key result is a characterization of the equilibrium steady-state marginal externality from vaccination. Across the continuum of market structures, the equilibrium steady-state marginal externality is non-monotonic in disease transmissibility: the marginal externality is zero if transmissibility is sufficiently small that the disease dies out in steady-state, peaks when the disease is just transmissible enough to survive in steady-state, and then weakly declines as transmissibility increases. Intuitively, this non-monotonicity arises because if transmissibility is very low then vaccinating an additional consumer does not provide other consumers with much protection since they would have been unlikely to contract the disease from that consumer; on the other hand, if transmissibility is very high then vaccinating an additional consumer does not provide other

[1] This paper predated the emergence of Sars-CoV-2, and thus does not focus on COVID-19. Our model is of the SI variety, while an SIR or possibly SIS model would be better suited for COVID-19. However, the last pandemic with impact at the scale of COVID-19 was HIV, and we can expect future pandemics that, like HIV, follow an SI dynamic. Work on adapting this paper's approach for COVID-19 is ongoing.

consumers with much protection since they are very likely to contract the disease from another source. Hence, our analysis suggests that marginal externalities from vaccination—and hence, governments' willingness to pay for additional vaccinations—are high for serious but rare diseases relative to diseases with lower individual burden but higher disease prevalence.

The intuition behind this non-monotonicity is most easily summarized when the vaccine is competitively produced and sold at cost. The marginal externality from vaccination is the difference between the marginal social benefit and the marginal private benefit. First consider the marginal social benefit. Our epidemiological framework is a standard but simple susceptible/infected model. A key feature of such models is that in steady-state, introducing an infected individual into a completely susceptible population generates exactly one new infection: if less than one, the disease would die out on its own; if more than one, disease prevalence would grow to a higher steady-state level. Given this structure of the model, each successfully vaccinated individual generates a unit reduction in steady-state disease prevalence. Intuitively, this arises because disease transmission is successful if an infected individual contacts a susceptible individual, and is unsuccessful if an infected individual contacts either another infected or a vaccinated individual. Because the steady-state rate of change of the infected population must equal zero, in steady-state each successfully vaccinated individual is a perfect substitute for an infected individual, implying that each successfully vaccinated individual generates a marginal social benefit of a unit reduction in steady-state disease prevalence.

Next, consider the marginal private benefit. Consumers purchase the vaccine if the cost is less than their marginal private benefit, which equals the product of the probability the vaccine is effective, flow harm from the disease conditional on infection, life expectancy, and the probability of infection. The last parameter—the probability of infection—depends on disease prevalence and transmissibility. This consumer decision rule generates a unique steady-state level of disease prevalence at which consumers are just indifferent about whether to purchase the vaccine. Hence, the (constant) vaccine cost pins down the marginal private benefit.

Combining these expressions for the marginal social benefit and marginal private benefit reveals that the marginal externality is non-monotonic in disease transmissibility. In particular, we obtain a counter-intuitive result: while an increase in transmissibility might be expected to always cause harm to society, if disease prevalence is in a range where some people purchase the vaccine and others free ride, then increases in transmissibility lead to a *reduction* in steady-state disease prevalence. Intuitively, increases in transmissibility lead to increased vaccination—as consumers

adjust to equate the cost and benefit of vaccination—and because the risk of infection is pinned down by the (constant) vaccine cost, a rise in transmissibility leads to a sufficient number of additional vaccinations to induce a decline in steady-state disease prevalence. This implies that while the social benefit of one additional vaccination is always a unit reduction in steady-state disease prevalence, for common diseases most of this benefit accrues to the vaccinated individual, while for rare diseases most of this benefit accrues to others in the form of externalities.

Building on this characterization of the marginal externality, for each market structure we derive the minimum per-dose subsidy required to attain the first best. Crude calibrations suggest optimal subsidies may be very large relative to current levels in the case of competitively supplied products: for example, if a subsidy of $600 induced someone to be circumcised and this reduced the probability of HIV transmission by $\frac{2}{3}$, that level of subsidy would be "worth it" from a social perspective. In the case of a monopolist producer selling directly to consumers, per-dose subsidies are a less attractive instrument because the minimum required per-dose subsidy needed to achieve the first best can be enormous—for some diseases, fifteen times the cost of actually becoming infected with the disease. Given plausible values for the deadweight loss of taxation, such subsidies would be prohibitively costly; bulk purchase of the optimal quantity followed by subsidized distribution could achieve the optimal level of immunization more cheaply.

Our baseline model focuses on homogenous consumers. Although this set-up may be reasonable for diseases such as measles and influenza, a natural question is whether adding heterogeneity would reduce the level of the (surprisingly large) optimal subsidies which are predicted by our model. First, we argue that heterogeneity in costs of vaccination are unlikely to overturn this result, but would generate a large value of targeting in avoiding payment of subsidies to infra-marginal consumers. Second, we argue that heterogeneity in disease transmissibility—for example, malaria being much more transmissible in Africa (for environmental reasons) than in the United States—would generate a large value of setting different subsidies for epidemiologically distinct subgroups. In an appendix, we show more formally that our key results are robust to incorporating at least some types of heterogeneity. Although a full analysis of all potential sources of heterogeneity is beyond the scope of this paper, our key results seem unlikely to be overturned by heterogeneity, but some types of heterogeneity would generate a large value of targeting subsidies.

While our epidemiological framework is a standard model, our novel contribution relative to the epidemiological literature is incorporating this model into a welfare economics framework which

allows us to analyze optimal policy.[2] Economists have long observed that vaccines may provide positive externalities that could affect consumers' and firms' decisions (see, among others, Brito, Sheshinski, and Intrilligator 1991; Boulier 2010; Chen and Toxvaerd 2014; Francis 1997; Geoffard and Philipson 1997; Gersovitz 2003; Gersovitz and Hammer 2004, 2005).[3] Boulier, Datta, and Goldfarb (2007) use a standard epidemiological model alone (i.e., neither interacted with consumer decisions nor a supply-side model of firm behavior) to examine properties of vaccination externalities that arise solely due to epidemiological concerns. Geoffard and Philipson (1997) use an epidemiological model similar to ours to show that a vaccine producer with market power will not choose to eradicate the disease in the steady-state. Galeotti and Rogers (2013) model vaccination choices in a heterogenous population, and consider the effect of network structures in determining optimal vaccine allocation.[4] Recent work has also estimated vaccine externalities empirically (Bethune and Korinek 2020; Cook et al. 2009; Ward 2014). Relative to the previous literature, our contribution is to incorporate a model of firm behavior and explicitly characterize equilibrium solutions for both positive and normative outcomes—specifically, externalities and optimal subsidies—in terms of estimable parameters.

The remainder of the paper is structured as follows. Section 2 outlines our epidemiological model and economic assumptions, which Section 3 then integrates into an economic model of vaccine supply and demand, analyzing the cases of both perfect competition and monopoly market structures. Section 4 characterizes externalities and optimal subsidies in our model. Section 5 discusses the implications of various forms of heterogeneity. Section 6 examines a firm's incentives to develop a vaccine relative to a drug that is identical except for the lack of the vaccine's externality. The final section concludes. A series of appendices provide omitted proofs and a number of extensions of the model.

[2]Much of the epidemiological literature focuses on characterizing the conditions under which it would be feasible to eradicate a disease, whereas our framework allows us to analyze optimal policy even though eradication is impossible in our model. Althouse, Bergstrom, and Bergstrom (2010) also consider welfare analysis of vaccination, calibrating a simple model for four prominent diseases to estimate optimal subsidies under perfect competition and perfectly effective vaccination. Our work complements theirs, allowing for imperfect vaccines, including a supply-side model of firm behavior, and generating comparative statics which allow theoretical insights into how epidemiological and economic parameters impact market outcomes and optimal policy.

[3]Recent work in behavioral epidemiology implicitly incorporates externalities, considering, for example, game-theoretic analyses of decisions around whether to vaccinate or to free ride on herd immunity. (Funk et al. 2010; Manfredi and D'Onofrio 2013)

[4]Mecholan (2007) provides some analysis of treatments (conditional on infection) for communicable diseases in the context of a monopoly manufacturer, but provides no analytical results, instead focusing on numerical simulations, primarily related to issues of drug resistance.

2. Model

In this section, we outline a standard epidemiological model (Section 2.1) and our economic assumptions and welfare framework (Section 2.2) which, in later sections, we will use to solve for steady-state equilibrium under alternative market structures. For brevity, we focus the model on the example of a vaccine market, but the analysis is applicable to any preventative technology.

2.1. Epidemiological Component

Consider a standard epidemiological model (Bailey 1975, Anderson and May 1991) of a non-fatal disease in which consumers are born and die at rate $\gamma > 0$. Since the disease is non-fatal, both the infected and uninfected have the same death rate (γ) and life expectancy ($1/\gamma$). Because the birth rate equals the death rate, the size of the population is constant over time. Normalize the population mass to unity. Let $\theta \in (0,1)$ denote vaccine efficacy—i.e., the probability that a vaccinated consumer is successfully immunized against the disease.[5] Assume that if the initial dose is not effective for a consumer, further doses will not be effective either.

Let $V(t)$ be the proportion of the population that is successfully vaccinated at time t, $I(t)$ the proportion of the population that is infected, and $S(t)$ the remaining susceptible proportion. Since these categories are exhaustive and mutually exclusive and since the mass of consumers is normalized to unity, we have

$$V(t)+I(t)+S(t) = 1. \tag{1}$$

Note that $V(t)$, $I(t)$, and $S(t)$ can be interpreted as either proportions or masses.

Letting $Q(t)$ be the mass of newborns who are vaccinated, the mass of newborns who are successfully vaccinated is $\theta Q(t)$. Given the Poisson structure of the model, and hence the stationarity of consumers' life cycles, assuming vaccines are purchased only by newborns is without loss of generality; we could equivalently have assumed that the vaccine is purchased by any arbitrary subset of susceptible consumers. For now we will take $Q(t)$ as given; we will later solve for its equilibrium value using the economic model and substitute the value back into the epidemiological model. The rate of change of the successfully vaccinated population is

$$\dot{V}(t) = \theta Q(t) - \gamma V(t), \tag{2}$$

[5] We take θ to be an exogenous parameter. Our main conclusions continue to hold if θ is endogenized by having the marginal cost be an increasing function $c(\theta)$ of efficacy.

the number of new successful vaccinations minus the number of vaccinated individuals who die.

Susceptible consumers contract the disease from infected consumers at rate $\beta I(t)$. Here, β embodies both the rate of contact with infected consumers and the rate at which those contacts lead to infection. The infection rate is linear in the number of infected individuals, $I(t)$. The rate of change of the infected population is thus

$$\dot{I}(t) = \beta I(t)S(t) - \gamma I(t), \tag{3}$$

the number of new infections from contact between the susceptible and infected populations minus the number of infected individuals who die.

To simplify the below analysis, we now introduce a convenient change of variables. A key parameter in epidemiology is the basic reproductive number of a disease, denoted R_0, which is defined as the average number of secondary infections an infected person would cause over her lifetime if she were introduced into an entirely susceptible population with no vaccinations (see, *e.g.*, Anderson and May 1991). If $R_0 < 1$, the disease eventually dies out in the population even without vaccination; if $R_0 > 1$, disease prevalence asymptotes toward an upper bound.

In the model, a single infected person spreads secondary infections at the rate β in a fully susceptible population, creating $R_0 = \beta/\gamma$ secondary infections over a lifespan of $1/\gamma$. Since we will treat γ as fixed throughout this paper, we can measure infectiousness relative to lifespan. Substituting the change of variables $\beta = \gamma R_0$ in equation (3) yields the equivalent expression:

$$\dot{I}(t) = \gamma R_0 I(t)S(t) - \gamma I(t). \tag{4}$$

In our comparative-statics exercises, we will take R_0 to be an exogenous parameter, capturing disease transmissibility. Table 1 presents estimates of R_0 from Anderson and May (1991). For conciseness, we present only estimates from England and Wales, the region for which Anderson and May (1991) have the most complete set of R_0 estimates. R_0 ranges considerably across diseases, from 2-5 for HIV at the low end to 16–18 for measles and pertussis at the high end. In addition to varying across disease, estimates of R_0 can also vary widely across time and region.

Let variables without time t arguments (Q, V, I, S, etc.) denote steady-state values. Taking Q as given for now (we will solve for the equilibrium value later), the steady-state values of V, I, and S can be found by substituting $\dot{V}(t) = \dot{I}(t) = 0$ in equations (2) and (4) and then forming a system

Table 1: Estimates of basic reproductive rate, R_0

Infection	Time period	R_0
HIV (Type I)	1981–85	2–5
Rubella	1960–70	6–7
Chicken pox	1944–68	10–12
Mumps	1960–80	11–14
Measles	1950–68	16–18
Pertussis	1944–78	16–18

Source: Estimates for England and Wales from Table 4.1 of Anderson and May (1991). *Notes:* The HIV (Type 1) estimate is for the subpopulation of men who have sex with men (MSM).

of equations with (1). The unique stable solution is[6]

$$I = \max\left\{0, 1-\frac{\theta Q}{\gamma}-\frac{1}{R_0}\right\} \tag{5}$$

$$V = \frac{\theta Q}{\gamma} \tag{6}$$

$$S = \min\left\{1-\frac{\theta Q}{\gamma}, \frac{1}{R_0}\right\}. \tag{7}$$

2.2. Economic Component

We consider the case of homogeneous,[7] risk-neutral consumers, and (to avoid computing transition paths) focus on the steady-state in the limiting case without discounting. Let $c > 0$ be the marginal cost of manufacturing and administering the vaccine. Let $h > 0$ denote the flow benefit experienced by a consumer at each instant she is healthy, a benefit which disappears once infected; thus h can also be interpreted as the flow harm from the disease.

We define the total health value as $HV^* = h(1-I^*)$, the product of flow health benefit to an

[6] A solution involving $I = 0$ always exists, but it is unstable when $R_0 > \gamma/(\gamma-\theta q)$.

[7] We discuss heterogeneity more in later sections. Appendix C shows that our central results extend to the case of consumers with heterogeneous valuation of health, for example, due to differences in income. Section 5 discusses the implications of other sources of heterogeneity.

individual h and the proportion of healthy individuals at any instant in steady-state equilibrium $1-I^*$.[8] We let W^* denote social welfare, the difference between total health value and total vaccine production costs: $W^* = HV^* - cQ^*$.

Given the Poisson structure of the model and the absence of discounting, a consumer's maximum willingness to pay for a vaccine in the steady-state—i.e., her marginal private benefit *MPB*—can be written

$$MPB = \frac{\theta h}{\gamma}\left(\frac{IR_0}{1+IR_0}\right). \tag{8}$$

MPB is the product of four factors: (i) vaccine efficacy, θ; (ii) flow harm from the disease conditional on contracting the disease, h, (iii) life expectancy, $1/\gamma$; and (iv) the probability of infection over the lifespan of an unvaccinated consumer, given by

$$\int_0^\infty \gamma R_0 I e^{-(\gamma+\gamma R_0 I)t} dt = \frac{IR_0}{1+IR_0}. \tag{9}$$

Equation (9) integrates over each instant t the product of two terms: $\gamma R_0 I$, which the argument preceding equation (4) showed is equal to the hazard of a susceptible person becoming infected, and $e^{-(\gamma+\gamma R_0 I)t}$, which is the probability that the person "survived as susceptible" to time t, neither becoming infected nor dying of natural causes.

Denote the marginal health value (in the sense of the marginal social benefit) by $MHV^* = dHV^*/dQ^*$. Taken together, $MX^* = MHV^* - MPB^*$ then denotes the marginal externality generated by a vaccine dose.

3. Market Equilibrium

This section provides an analysis of the steady-state market equilibrium in two cases—perfectly competitive supply and monopoly supply. A more general Cournot competition model nests each of these extremes; Cournot results are provided in Appendix B. The case of perfect competition is more relevant for "commodity" vaccines and for other preventative commodities—for example for water treatment, condoms, and mosquito nets. The monopoly case is more relevant for manufacturers of on-patent vaccines; although market power has not been the focus of prior literature,

[8] HV^* is a measure of the population's overall health rather than the vaccine's contribution to health per se. When $R_0 < 1$, the population is completely healthy, but this happens even without any vaccinations. The marginal benefit measures derived below will more directly measure the vaccine's per se contribution.

this case is important for understanding research and development (R&D) incentives.[9] Section 3.1 derives consumer demand, Section 3.2 considers supply under competition and monopoly, and Section 3.3 solves for the steady-state equilibrium. Section 3.4 provides an illustrative calibration.

Notationally, let P denote the market price in the steady-state, and let Q and Π denote market quantity and industry profit (also firm quantity and profit in the monopoly case), both measured as flows each instant in the steady-state. Equilibrium values of steady-state variables will be distinguished with asterisks: P^*, Q^*, Π^*, I^*, and so forth. To eliminate trivial cases, we assume throughout the remainder of the paper that the disease is harmful enough that the social benefit of prevention exceeds the cost:

$$\theta h/\gamma > c. \tag{10}$$

We will show (see equation (17)) that $\theta h/\gamma$ is the marginal health value—i.e., the social benefit—from an additional dose of the vaccine. Thus, if (10) were violated, the vaccine would be so costly that none would be consumed in the first best, let alone in equilibrium in which externalities lead to underconsumption.

3.1. Demand

In deciding whether or not to buy the vaccine, consumers compare the private benefit to the vaccine's price. If $P < MPB$ consumers strictly prefer to buy the vaccine, if $P > MPB$ they strictly prefer not to buy it, and if $P = MPB$ they are indifferent. MPB is increasing in disease prevalence I, indicating that consumers are willing to pay more for the vaccine the more prevalent is the disease, and thus the greater the chance of being infected, all else equal.

Setting $P = MPB$ in equation (8) and solving for I yields the cutoff prevalence rate

$$\hat{I}(P) = \frac{\gamma P}{(\theta h - \gamma P)R_0} \tag{11}$$

such that consumers are indifferent between buying the vaccine and not if $I = \hat{I}(P)$, strictly prefer

[9] Compared to the number of firms that manufacture drugs, the number of vaccine manufacturers is very small. One reason for limited entry into vaccine manufacturing is the difficulty in replicating the production process, which is much less standardized than is manufacturing for drugs. Whereas patents are the key barrier to entry for drugs and production itself is not difficult to reverse engineer, this is not the case for vaccines. In 2002, only four firms were producing almost all of the recommended childhood vaccines for the U.S. market (GAO 2002). Time-series data on vaccine entry paints a similar picture (see Berndt et al. 2007). After the approval of the initial vaccine for hepatitis B and varicella in the 1980s, a second firm did not enter the market until more than eight years later. In other cases (such as hepatitis A and haemophilus influenzae type b) in which subsequent entry happened quickly after the first vaccine was approved, still only two or three firms ended up entering the market in each case.

to buy if $I > \hat{I}(P)$, and strictly prefer not to buy if $I < \hat{I}(P)$. This yields the demand relation

$$Q = \begin{cases} 0 & I < \hat{I}(P) \\ [0,\gamma] & I = \hat{I}(P) \\ \gamma & I > \hat{I}(P). \end{cases} \tag{12}$$

3.2. Supply

In the perfect competition case, vaccine supply is perfectly elastic at price $P = c$.

Figure 1 illustrates the monopolist's problem. The box contains the set of feasible values of vaccine quantity Q and disease prevalence I. Q is bounded above by γ (all newborns vaccinated) and below by zero; I bounded above by $1-1/R_0$ (the steady-state infection rate from equation (5) with no vaccinations) and below by zero. The equilibrium simultaneously satisfies the epidemiological relation (5), plotted as the dotted line ER, and the demand relation (12), plotted as the solid line DR. The monopolist controls the height of the horizontal portion of the demand relation via its choice of price P, with higher values of P shifting the horizontal portion up. That is, the monopolist chooses the horizontal position of the demand relation to generate the most profitable intersection point.

In order to write out the constrained maximization problem that the monopolist solves to choose her optimal price and quantity, it is helpful to derive an expression for reduced form demand $D(P)$. Here, "reduced form" refers to the fact that demand $D(P)$ is expressed as a function of P alone, rather than as a function of both endogenous variables P and I as in the structural demand relation (equation (12)). $D(P)$ can be derived as the intersection of the epidemiological relation (equation (5)) and the demand relation (equation (12)): substituting the expression for I from equation (5) into equation (11) and rearranging yields the following expression

$$\tilde{D}(P) = \frac{\gamma}{\theta}\left[1 - \frac{\theta h}{(\theta h - \gamma P)R_0}\right]. \tag{13}$$

To ensure that demand is non-negative, we write reduced-form demand as $D(P) = \max\{0, \tilde{D}(P)\}$. The monopoly's optimal price P^* solves the constrained maximization problem

$$\max_P (P-c)D(P) \qquad \text{subject to} \quad D(P) \leq \gamma, \tag{14}$$

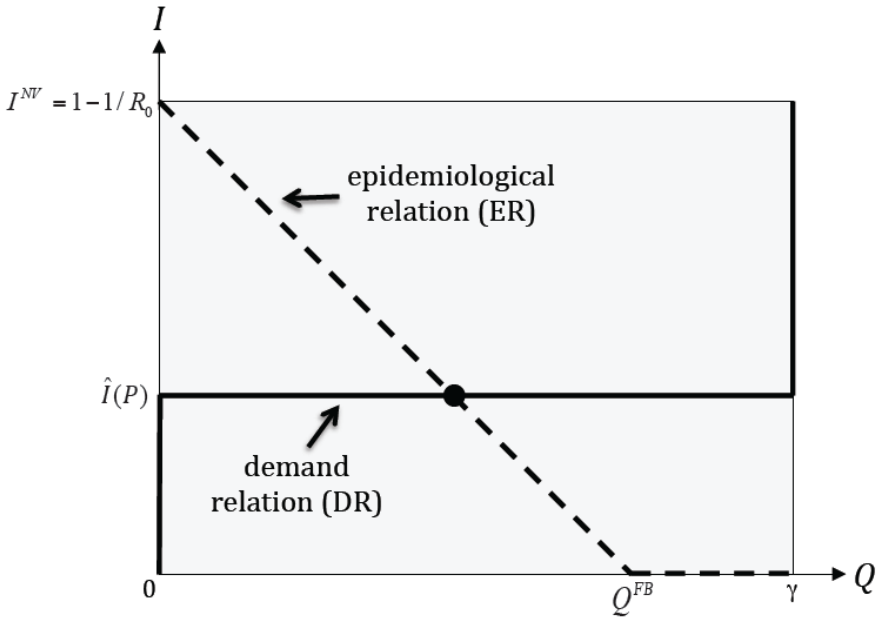

Figure 1: Monopolist's problem: Feasible vaccine quantity (Q)–disease prevalence (I) pairs. *Notes:* The horizontal axis represents vaccine quantity Q and vertical axis disease prevalence I. Dotted line plots the epidemiological relation (ER), and solid line plots the demand relation (DR). I^{NV} denotes the infection level in the absence of a vaccine; on the x-axis, Q^{FB} denotes the first-best vaccine quantity that would be chosen by a social planner.

where the constraint ensures that vaccine sales cannot exceed the number of consumers (the flow of γ newborns in the steady-state).

3.3. Equilibrium

For concision, it will be useful to define a new parameter ρ:

$$\rho = \frac{\theta h}{\theta h - \gamma c}. \tag{15}$$

Dividing through by γ, the numerator can be interpreted as the benefit of taking the vaccine if infection were certain in the absence of vaccination, and the denominator can be interpreted as this benefit less the cost of the vaccine. Note that if $c = 0$, $\rho = 1$; at the other extreme, as c approaches the level at it which it would be socially inefficient to consume the vaccine even if infection were automatic ($c \uparrow \theta h/\gamma$), ρ becomes arbitrarily large.

Intuition for the closed form solutions for the integrated economic and epidemiological model

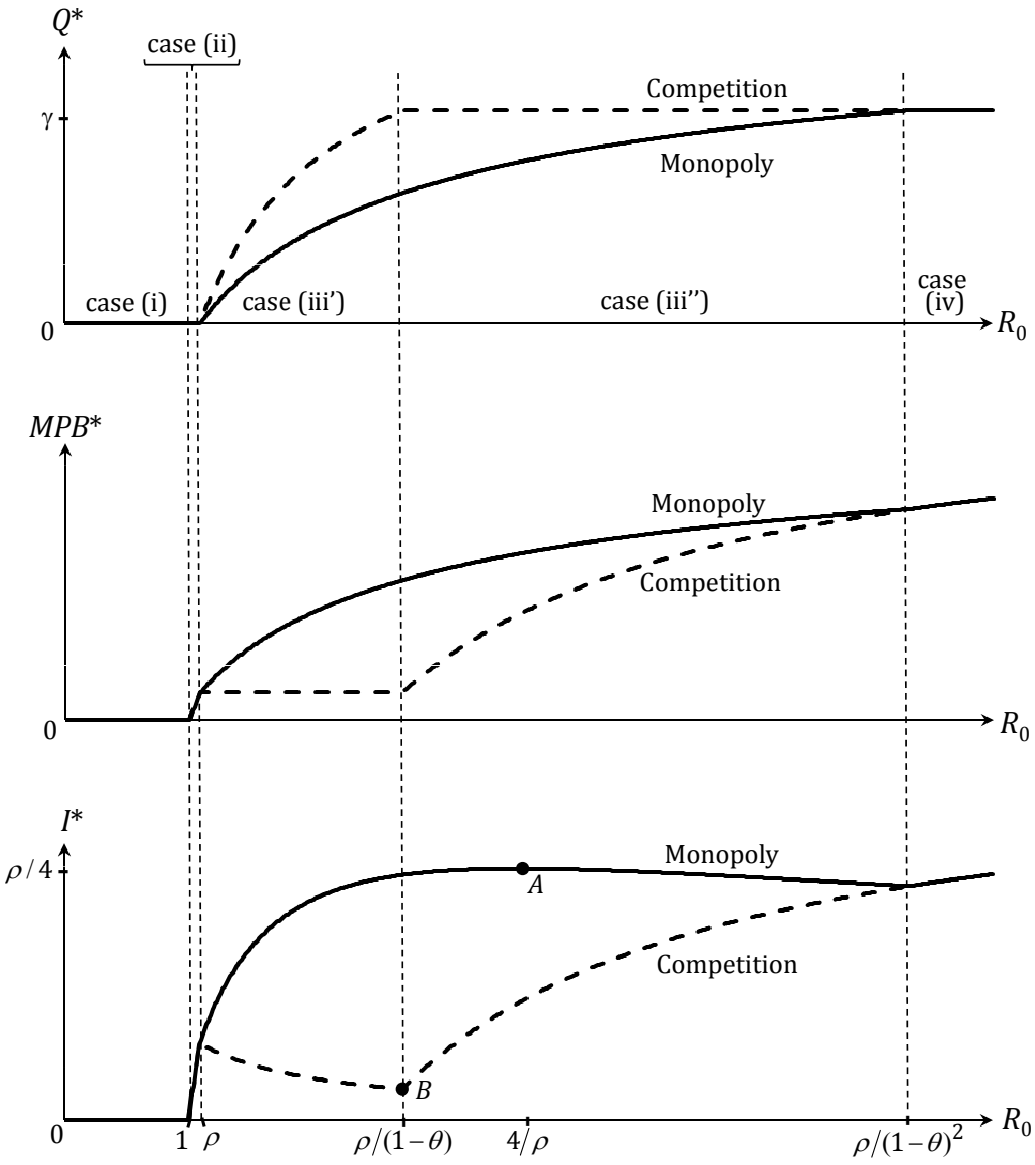

Figure 2: Equilibrium variables as functions of R_0

can be gained from examining Figure 2.[10] The top panel illustrates the equilibrium quantity of vaccine, Q^*; the middle panel illustrates the equilibrium marginal private benefit, MPB^*; and the bottom panel illustrates the equilibrium infection level, I^*. The dotted line represents the perfect competition case, and the solid curve represents the monopoly case.

The equilibrium depends on the level of transmission, R_0: as R_0 increases from 0, four cases arise in turn ((i)-(iv)). For very low levels of R_0 (cases (i) and (ii)), even in the no-vaccine steady-state, the private benefit from a unit of vaccine is smaller than the marginal cost of producing it, and

[10]Note that Figure 2 is drawn for parameter values such that all possible regions are illustrated. For $\theta = 1$, for example, case (iv) will not exist.

thus no vaccine is sold ($Q^* = 0$). While not very interesting from the perspective of our analysis, we briefly characterize these two cases for completeness. In case (i), R_0 is sufficiently low that the disease disappears in steady-state even with no vaccinations; I^* and MPB^* are zero. In case (ii), R_0 is high enough that the disease is not eradicated in steady-state, but still low enough to result in a steady-state disease prevalence such that it is not privately optimal to be vaccinated even if the vaccine is sold at cost. Since consumers are not purchasing the vaccine in this range, infection risk I^* unambiguously grows as R_0 increases, implying that MPB^* increases in R_0 throughout case (ii).

As R_0 increases, we eventually reach a point at which $MPC^* = c$, marking the lower bound on our primary case of interest: case (iii). Case (iii) encompasses all R_0 values sufficiently high to generate some vaccine consumption, but not so high as to lead to full vaccination of all newborns.

Consider first the case of perfect competition. As disease transmissibility R_0 rises, the probability of infection rises for any given level of vaccination. Thus vaccine demand Q^* increases. Thus in this range an increase in R_0 has both a direct effect (higher I for any given level of vaccination) and an indirect effect (increased vaccine demand) on disease prevalence. For any interior steady-state equilibrium, in which some people are vaccinated and some are not, we must have $MPB^* = c$. Thus in equilibrium the risk of infection is pinned down by c, so I^* must decrease to offset increases in R_0. We see, counterintuitively, that as R_0 increases its indirect effect on I^* must outweigh its direct effect so that disease prevalence is *increasing* in transmissibility.

Now consider the monopoly case. Note that weakly fewer people are treated than under competition, because selling the vaccine to one consumer reduces others' willingness to pay. While it remains the case that vaccine demand Q^* increases with R_0, the monopolist restricts the increase in Q^* so that MPB^* grows with the infectiveness of the disease. An interior equilibrium is established such as that shown in Figure 1, in which the cost of increased disease risk from remaining unvaccinated exactly equals the vaccine's price P^*. Because fewer consumers are vaccinated than under competition, MPB^* and I^* are higher.

Under the assumption that the vaccine is not perfectly effective, R_0 eventually reaches a point at which the disease is so transmissible that vaccination hits a corner solution with all γ newborns buying the vaccine at each instant. Note that MPB^* still rises with R_0 above this point because the chance that vaccinated but unsuccessfully immunized consumers contract the disease increases. This case occurs at a higher level of transmissibility under monopoly than under competition, since the monopolist raising P^* slows the increase in Q^*. This case does not arise at all with a perfectly effective vaccine ($\theta = 1$) because the marginal consumer would never be willing to pay a

positive price in the steady-state if everyone else were using a perfectly effective vaccine.

For ease of presentation, we will abuse notation as follows. We let case (iv) denote the case in which all newborns are vaccinated, but as noted, this occurs for a larger set of R_0 values under competition than under monopoly. Thus when referring to the monopoly case, we let case (iv) cover the range where $R_0 \geq \rho/(1-\theta)^2$. When referring to the perfect competition case, case (iv) covers the range where $R_0 > \rho/(1-\theta)$. When considering both the monopoly and competition cases, as in Figure 2 or Table 2, we let (iii') denote the case in which some but not all newborns are vaccinated under both competition and monopoly, (iii") denote the case in which vaccination is full under competition but not under monopoly, and (iv) denote the case in which vaccination is full under both market structures.[11]

Proposition 1 formalizes these results. The proofs of this and all subsequent propositions and lemmas are provided in Appendix A.

Proposition 1. *The steady-state equilibrium falls into one of the cases in Table 2, depending on the market structure (monopoly or perfect competition) and on the level of transmission* R_0.

As explained above, under perfect competition, equilibrium disease prevalence I^* is decreasing in R_0 in case (iii) and increasing in case (iv). Since no one is vaccinated in case (ii), I^* is increasing in R_0 in this case as well. Thus, as can be seen in the bottom panel of Figure 2, I^* reaches a local maximum at the infectiousness level $R_0 = \rho$ dividing cases (ii) and (iii), and a local minimum at the level $R_0 = \rho/(1-\theta)$ dividing cases (iii) and (iv).

The comparative statics of I^* are not as clear-cut in the case of monopoly. In cases (ii) and (iv), I^* is equivalent to the perfect competition case, and is thus increasing in R_0. However, in case (iii), recall that the monopolist weakens the indirect effect of R_0 on I^* by raising prices. Thus in case (iii), the direct effect is initially stronger and I^* is increasing in R_0. For some parameter values (such as those used in Figure 2), I^* is decreasing for R_0 in the upper range of case (iii) and thus reaches a local maximum at the point labeled A. As Proposition 2 indicates, I^* is nonmonotonic in R_0 under monopoly if and only if $\rho > 2(1-\theta)$. This inequality is likely to be satisfied in practical situations: any vaccine that the social planner would use that is more than 50% effective satisfies $\rho > 2(1-\theta)$. For the parameters used in the figure, the local maximum attained by I^* in R_0 is not a global maximum. However, as Proposition 2 indicates, it will be a global maximum

[11] Since $P^* = MPB^*$ when the monopolist serves some consumers (cases (iii) and (iv)), the middle panel also illustrates how P^* varies with R_0. To save space, the graph for Π^* is not shown; it resembles the top two panels of Figure 2.

Table 2: Equilibrium variables as functions of R_0

	Interval for R_0				
Variable	$[0,1]$ case (i)	$(1,\rho)$ case (ii)	$\left[\rho, \frac{\rho}{1-\theta}\right]$ case (iii′)	$\left(\frac{\rho}{1-\theta}, \frac{\rho}{(1-\theta)^2}\right)$ case (iii″)	$\left[\frac{\rho}{(1-\theta)^2}, \infty\right)$ case (iv)
Panel A: Monopoly					
Q_m^*	0	0	$\frac{\gamma}{\theta}\left(1-\sqrt{\frac{\rho}{R_0}}\right)$	same as (iii′)	γ
I_m^*	0	$1-\frac{1}{R_0}$	$\sqrt{\frac{\rho}{R_0}}-\frac{1}{R_0}$	same as (iii′)	$1-\theta-\frac{1}{R_0}$
MPB_m^*	0	$\frac{\theta h}{\gamma}\left(1-\frac{1}{R_0}\right)$	$\frac{\theta h}{\gamma}\left(1-\frac{1}{\sqrt{\rho R_0}}\right)$	same as (iii′)	$\frac{\theta h}{\gamma}\left[1-\frac{1}{(1-\theta)R_0}\right]$
Π_m^*	0	0	$\frac{h}{\rho}\left(1-\sqrt{\frac{\rho}{R_0}}\right)^2$	same as (iii′)	$\theta h\left[\frac{1}{\rho}-\frac{1}{(1-\theta)R_0}\right]$
P_m^*	$(0,\infty)$	$[c,\infty)$	same as MPB_m^*	same as (iii′)	same as MPB_m^*
Panel B: Competition					
Q_c^*	0	0	$\frac{\gamma}{\theta}\left(1-\frac{\rho}{R_0}\right)$	γ	same as (iii″)
I_c^*	0	$1-\frac{1}{R_0}$	$\frac{\gamma c\rho}{\theta h R_0}$	$1-\theta-\frac{1}{R_0}$	same as (iii″)
MPB_c^*	0	$\frac{\theta h}{\gamma}\left(1-\frac{1}{R_0}\right)$	c	$\frac{\theta h}{\gamma}\left[1-\frac{1}{(1-\theta)R_0}\right]$	same as (iii″)
Π_c^*	0	0	0	0	0
P_c^*	c	c	c	c	c

Notes: For entries that are intervals, variable can take on any value in interval.

for $\rho \geq 4(1-\theta)$, a condition that is satisfied for any vaccine that the social planner would use that is more than 75% effective. In practice, agencies such as the U.S. Food and Drug Administration (FDA) do not typically license vaccines with efficacy below 80% (Brennan 2009). Thus, for realistic parameter values we have a surprising result: for a vaccine sold directly to consumers by a monopolist, the disease burden is greatest for diseases with moderate rather than the highest degree of infectiousness as indexed by R_0. The nonmontonicity of I^* can be significant: the value of I^* at the interior maximum at A can be arbitrarily higher in percentage terms than the limiting value of I^* as $R_0 \uparrow \infty$. As discussed below, this implies distortions associated with monopoly are particularly great for monopolists selling directly to consumers for intermediate values of transmissibility.

Proposition 2 formalizes the comparative statics results.

Proposition 2. Comparative statics with respect to R_0 under perfect competition and monopoly are as follows:

(i) Q^* is weakly increasing in R_0 under both perfect competition and monopoly.

(ii) MPB^* is weakly increasing in R_0 under both perfect competition and monopoly.

(iii) Π^* is nondecreasing in R_0 ($\Pi^* = 0$ under perfect competition).

(iv) $P^* = c$ under perfect competition.P^* is nondecreasing in R_0 under monopoly for some series of equilibrium monopoly prices.

(v) Under perfect competition, I^* is nonmonotonic with a local maximum at $R_0 = \rho$; this local maximum is a global maximum if $c > h/2\gamma$.

(vi) Under monopoly, if $\rho \leq 2(1-\theta)$, then I^* is nondecreasing in R_0; otherwise I^* is nonmonotonic with a local maximum at $R_0 = 4/\rho$, and this local maximum is a global maximum if and only if $\rho \geq 4(1-\theta)$.

3.4. Calibration

To give some sense for the implications of Proposition 1 for existing vaccines, it is useful to present an illustrative calibration using one particular set of parameter values for vaccine efficacy θ and cost of the vaccine c relative to harm from the disease h. As discussed above, the efficacy of available vaccines in practice is extremely high; we consider the case where $\theta = 1$. Because the cost of c for existing vaccines is usually quite low, especially in comparison to the flow harm from the disease h, we assume that the magnitude of c relative to h is close to zero, implying that ρ is close to 1. Using these assumptions, we can generate simplified equilibrium approximations that are only functions

of R_0. To give further intuition, we also present these equilibrium approximations re-writing R_0 as $\frac{1}{1-I^{nv}}$, where I^{nv} is the equilibrium infection level in the absence of a vaccine; this alternative set of approximations connects these equilibrium expressions to a second, potentially directly observable piece of data—namely, disease prevalence in the absence of a vaccine.

In case (iii) for the monopoly market structure,we then have that the equilibrium quantity Q^* as a share of the birth cohort γ is

$$Q_m^*/\gamma = \left(1-\sqrt{\frac{1}{R_0}}\right) = \left(1-\sqrt{1-I^{nv}}\right)$$

equilibrium disease prevalence I^* is

$$I_m^* = \sqrt{\frac{1}{R_0}} - \frac{1}{R_0} = \sqrt{1-I^{nv}} - 1 + I^{nv}$$

price P^* relative to the expected harm from the disease conditional on infection h/γ is

$$P_m^*/\frac{h}{\gamma} = \left(1-\frac{1}{\sqrt{R_0}}\right) = \left(1-\frac{1}{\sqrt{\frac{1}{1-I^{nv}}}}\right) = MPB^*/\frac{h}{\gamma}$$

and the flow of equilibrium profits Π^* relative to the flow social harm from the disease per fraction of each cohort infected, h,is

$$\Pi_m^*/h = \left(1-\sqrt{\frac{1}{R_0}}\right)^2 = \left(1-\sqrt{1-I^{nv}}\right)^2.$$

Although the assumptions that $\theta \approx 1$ and that $\frac{c}{h} \approx 0$ are important simplifications, they allow our key equilibrium expressions to be approximated as very simple functions of only one estimable parameter: R_0–or, alternatively, steady-state prevalence in the absence of a vaccine I^{nv}. Thus, given empirical estimates from the epidemiology literature of how R_0 varies across diseases, or current prevalence levels for disease without vaccines, these expressions provide very clear implications for the equilibrium infection level vaccine prices, quantities, and profits under a system in which a monopolist sells directly to consumers. For example, if one takes the estimated R_0 of 4 for HIV among men who have sex with men (MSM) in England and Wales from Table 1 seriously, this suggests that a monopolist selling to this market would price a vaccine at half the cost of being infected with the disease conditional on infection; that at this price one half of consumers would

choose to be vaccinated and the other half would free ride on others' decisions to be immunized; and that half of the free riders would become infected, so the overall prevalence rate would be 1/4. Profits would be one quarter the flow harm from the disease. Note that in the absence of a vaccine, with $R_0 = 4$, three quarters of the population would be infected; if a costless vaccine were developed the social planner would immunize enough people to reduce prevalence to zero, and moreover, this would also be the outcome under competition. So this implies that the vaccine manufacturer captures 1/3 of the potential social surplus from a vaccine, consumers capture 1/3 (since they face a 50% rather than 75% risk of infection), and 1/3 is lost due to monopoly distortions.

It is worth noting that the model implies that under a system in which a monopolist sells directly to consumers, I^* would peak at a value of $R_0 = 4$, as shown in Figure 2 (given our assumption here that $\rho = 1$). This implies that the equilibrium infection level I^* would be highest for diseases with estimated R_0 values close to 4, such as the example of HIV in Table 1. The share of consumers vaccinated in equilibrium approaches 1 as R_0 increases, and declines towards zero as R_0 approaches 1 from above.

The simplified expression for the equilibrium price as a share of the cost of certain infection with the disease clarifies that this share would be around 3/4 for diseases with values of R_0 around 16 such as measles, and approaches 0 as R_0 becomes closer to 1. Chowell et al. (2003) estimated values of R_0 for SARS in Singapore of 1.1, implying the equilibrium price as a share of the cost of being infected with SARS would be close to zero.

4. Normative Analysis

For the normative analysis, we start with a measure of total social welfare (Section 4.1) and then derive the marginal externality from additional vaccinations (Section 4.2) and the minimum government expenditure implementing the first-best level of vaccination (Section 4.3). Section 4.4 provides an illustrative calibration.

Recall that $HV^* = h(1-I^*)$ is the total health value, i.e., the product of flow health benefit to an individual h and the proportion of healthy individuals at any instant in steady-state equilibrium $1-I^*$. The following result will be useful in the subsequent analysis.

Lemma 1. Under any market structure with the property that $P^* > 0$ when $Q^* > 0$, the total health

value in the steady-state equilibrium can be expressed as the following function of Q^*:

$$HV^* = \begin{cases} h & \text{if } R_0 \leq 1 \\ h\left[\dfrac{\theta Q^*}{\gamma} + \dfrac{1}{R_0}\right] & \text{if } R_0 > 1. \end{cases} \tag{16}$$

Lemma 1 follows almost immediately from the epidemiological relation (5). The one technical detail in the proof is to show that some cases can be dispensed with because they never arise in equilibrium. In particular, under any market structure in which vaccine consumers pay a positive price, vaccination cannot eradicate the disease because the marginal consumer would not pay a positive price in the absence of infection risk. This confirms that the result of Geoffard and Philipson (1997) holds in our model. For such market structures, we can rule out cases in which $R_0 > 1$ yet $I^* = 0$. The lemma applies to a broad class of market structures as the price generally will be positive when $c > 0$ unless the government provides a subsidy or intervenes in some other way.

4.1. Welfare

Recall that W^* denotes social welfare, the difference between total health value and total production costs: $W^* = HV^* - cQ^*$. Combining this equation with Lemma 1 leads to a characterization of W^*, shown in Proposition 3. Before presenting these formal solutions, the intuition behind the results can be gained from examining the top panel of Figure 3, which gives a graphical representation of Propositions 3 and 4.

The most striking feature of the graph of W^* in the monopoly case is that it increases in R_0 for a range of values in case (iii). One might expect that an increase in the infectiousness of a disease as indexed by R_0 would always strictly harm society. While an increase in infectiousness of the disease as indexed by R_0 causes a direct harm to society, an offsetting effect in case (iii) is to induce more people to become vaccinated, helping counteract the problem of underconsumption in the presence of a positive externality. Under monopoly, not only do consumers fail to consider the external benefit their vaccination provides other consumers, but on top of that the monopolist places negative value on consumption to the extent it reduces others' willingness to pay for a vaccine. Mitigation of this severe underconsumption problem by increasing R_0 provides such a large benefit that it swamps the direct harm from an increase in R_0, leading to an increase in social welfare. The fact that a higher R_0 can actually be socially beneficial in the case of a monopolist can be interpreted as implying that monopoly distortions are particularly harmful at moderate levels of infection.

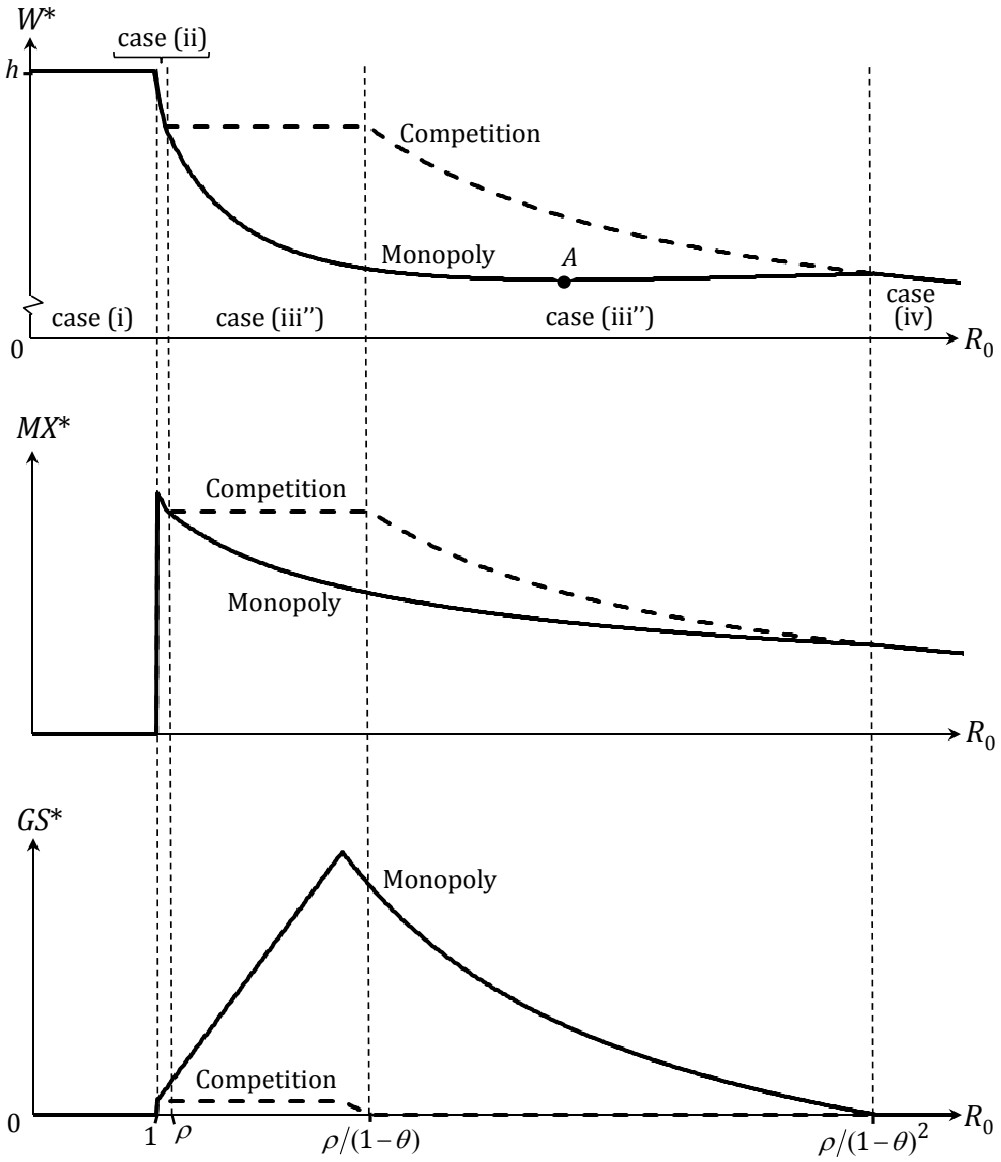

Figure 3: Normative variables as functions of R_0

Note first that because we assume consumers are homogeneous, we shut down the standard channel for welfare loss from monopoly. Nonetheless, W^* is higher under competition relative to monopoly: unlike the monopoly case, the only source of distortion under perfect competition is the vaccine externality, in which case welfare never increases with the infectiousness of the disease. W^* is weakly decreasing in R_0 in all cases under competition. Taken together, these results suggest market power creates little welfare loss for low values of R_0 or for high enough values of R_0 with an imperfect vaccine, and is most socially costly for intermediate values of R_0 or perfectly effective vaccines. Proposition 3 formalizes these results.

Table 3: Equilibrium welfare as a function of R_0

	Interval for R_0				
Variable	$[0,1]$ case (i)	$(1,\rho)$ case (ii)	$\left[\rho, \frac{\rho}{1-\theta}\right]$ case (iii′)	$\left(\frac{\rho}{1-\theta}, \frac{\rho}{(1-\theta)^2}\right)$ case (iii″)	$\left[\frac{\rho}{(1-\theta)^2}, \infty\right)$ case (iv)
Panel A: Monopoly					
W_m^*	h	$\frac{h}{R_0}$	$\frac{h}{\rho}\left(1-\sqrt{\frac{\rho}{R_0}}+\frac{\rho}{R_0}\right)$	same as (iii′)	$\theta h\left(\frac{1}{\rho}+\frac{1}{\theta R_0}\right)$
Panel B: Competition					
W_c^*	h	$\frac{h}{R_0}$	$\frac{h}{\rho}$	$\frac{\theta h}{\rho}+\frac{\rho}{R_0}$	same as (iii″)

Proposition 3. Steady-state equilibrium welfare falls into one of the cases in Table 3, depending on the market structure (monopoly or perfect competition) and on the level of transmission R_0. Under monopoly, W^* is nonmonotonic in R_0 with a local minimum at $R_0 = 4\rho$; this local minimum is a global minimum if $\theta > 3/4$.

4.2. Marginal Externality

Recall that we denoted the marginal health value (in the sense of the marginal social benefit) by $MHV^* = dHV^*/dQ^*$. Differentiating (16) yields

$$MHV^* = \begin{cases} 0 & \text{if } R_0 \leq 1 \\ \theta h/\gamma & \text{if } R_0 > 1. \end{cases} \tag{17}$$

Equation (17) has some interesting implications which deserve comment. In all nontrivial cases, the health value of an additional vaccine dose equals $\theta h/\gamma$, which is the expected benefit of using a drug (with the same efficacy, θ, as the vaccine) to treat an infected person. This might seem puzzling at first because the vaccinated individual was not certain to have contracted the disease, so it is not clear that the health value of administering a vaccine to someone who may not be infected should amount to the benefit from treating someone who has certainly been infected. The puzzle can be resolved by again considering the positive externality associated with the vaccine. Vacci-

nating an individual prevents a whole chain of possible infections in the unvaccinated population. The direct benefit to the vaccinated individual and the chain of expected external benefits turn out to sum exactly to $\theta h/\gamma$ in the model.

Another interesting implication of (17) is that the health value of an additional vaccine dose is a constant, independent of changes in the infectiveness of the disease as measured by R_0 (again, in all nontrivial cases). As we will see, the direct benefit on the vaccinated individual and the externality will each vary with R_0, but the sum of these effects is constant.

Recall that $MX^* = MHV^* - MPB^*$ is the marginal externality provided by a dose of the vaccine. Substituting the expression for MPB^* from Proposition 1 and for MHV^* from equation (17) into the preceding equation yields a characterization of MX^*. Before presenting these formal solutions in Proposition 4, the intuition behind the results can be gained from examining the middle panel of Figure 3. Because the disease dies out in the steady-state for $R_0 \leq 1$ (case (i)), the vaccine has no associated marginal externality in that case. MX^* jumps to its global supremum at $R_0 = 1$ and strictly declines in higher values of R_0; the implication is that diseases with a moderate rather than the highest level of transmissibility—in particular, those with R_0 near 1—may present the greatest case for public involvement. With lower levels of infectiveness, the disease dies out without a vaccine; with higher levels of infectiveness, consumers appropriate more of the health value of using the vaccine. Proposition 4 formalizes these results.

Proposition 4. The marginal externality in steady-state equilibrium falls into one of the cases in Table 4, depending on the market structure (monopoly or perfect competition) and on the level of transmission R_0. Under both monopoly and perfect competition, MX^* approaches a supremum as $R_0 \downarrow 1$.

4.3. Government Subsidies

The positive externality characterized in Section 4.2 will lead to too few consumers purchasing the vaccine relative to the social optimum. Consumers do not take into account the reduction in the spread of the disease to others from their own immunization. This naturally raises the question of whether government intervention can correct the market failure generated by this positive externality.

In this section, we show that a per-dose subsidy to a monopolist selling directly to individual consumers would—for many realistic parameter values—be a prohibitively expensive way for governments to reach the first best for any reasonable values of the cost of public funds. As we will

Table 4: Equilibrium marginal externality as a function of R_0

	Interval for R_0				
Variable	$[0,1]$ case (i)	$(1,\rho)$ case (ii)	$\left[\rho, \frac{\rho}{1-\theta}\right]$ case (iii′)	$\left(\frac{\rho}{1-\theta}, \frac{\rho}{(1-\theta)^2}\right)$ case (iii″)	$\left[\frac{\rho}{(1-\theta)^2}, \infty\right)$ case (iv)
Panel A: Monopoly					
MX_m^*	0	$\frac{\theta h}{\gamma R_0}$	$\frac{\theta h}{\gamma\sqrt{\rho R_0}}$	same as (iii′)	$\frac{\theta h}{\gamma(1-\theta)R_0}$
Panel B: Competition					
MX_c^*	0	$\frac{\theta h}{\gamma R_0}$	$\frac{\theta h}{\gamma\rho}$	$\frac{\theta h}{(1-\theta)\gamma R_0}$	same as (iii″)

see, there is always a sufficiently high subsidy such that the first best can be attained, where either all consumers purchase the vaccine or the steady-state infection rate is driven to zero. However, the subsidy would in some cases cost many times the value of avoiding a disease. This level of subsidy would be prohibitively expensive, and would also be problematic from a dynamic perspective in terms of creating too strong an incentive to develop a vaccine. This suggests that governments either use policy levers other than a per-dose subsidy, such as bulk purchases, or do not attempt to reach the first best.

Assume the benevolent government maximizes social welfare W^*. Let GS denote the per-dose subsidy to which the government commits at the outset of the game. We will adopt the accounting convention that the subsidy is paid to the firm. The subsidy has the same effect on the monopolist as a reduction in its marginal cost from c to $c-GS$.

Proposition 5 considers the the lowest possible subsidy attaining the first best. The bottom panel of Figure 3 shows a graphical version of how GS^* varies with R_0. While the graph of GS^* has the same general hump shape as the graph of MX^*, the two are not identical for several reasons. First, MX^* reflects society's maximum willingness to pay for an additional dose to be sold in equilibrium. Society may be able to get away with paying less than the full societal value for this dose, depending on the monopolist's markup over the subsidy. Second, MX^* represents the externality at one particular point: in equilibrium. The optimal subsidy moves the market to a

Table 5: Equilibrium lowest possible subsidy attaining the first best as a function of R_0

	Interval for R_0	
Variable	$\left(1, \frac{1}{1-\theta}\right]$ case (a)	$\left(\frac{1}{1-\theta}, \frac{\rho}{(1-\theta)^2}\right)$ case (b)
Panel A: Monopoly		
GS_m^*	$c+\frac{\theta h}{\gamma}$	$c+\frac{\theta h}{\gamma}\left[\frac{1}{(1-\theta)^2 R_0}-1\right]$
Panel B: Competition		
GS_c^*	c	$c-\frac{\theta h}{\gamma}\left[\frac{1}{(1-\theta)R_0}-\frac{1}{\rho}\right]$

discretely different point, presumably one involving greater vaccine sales, and GS^* depends on the externality level at this other point.[12] GS^* is lower under competition than monopoly. Proposition 5 formalizes these results.

Proposition 5. The minimum subsidy implementing the first best level of vaccine use under either market structure (monopoly or perfect competition) when vaccines are sold directly to consumers is $GS^* = 0$ in cases (i) and (iv). The other cases—which differ from the cases in Proposition 1—are summarized in Table 5. GS^* is hump-shaped in R_0. In the monopoly case, GS^* attains a maximum at $R_0 = 1/(1-\theta)$; in the competitive case, GS^* attains a maximum for all R_0 in the interval $(1, 1/(1-\theta)]$.

One implication is that for an efficacious vaccine sold by a monopolist, if R_0 is high then the extra subsidy required to attain the first best will be huge, many times the cost of becoming infected with the disease. For a high R_0, efficiency requires immunizing almost everyone. For a monopolist to be induced to drop its price to consumers enough to pick up the last few consumers, it will need to be compensated for the reduced revenue on existing consumers with a very high subsidy. For a disease like measles, vaccine efficacy is close to 1, and R_0 is perhaps 16. This implies that the minimum extra government subsidy to reach the first best would be 15 times the cost of certain infection with measles, probably thousands of dollars per-dose. With any deadweight loss

[12] Note that MX^* reflects the externality not just on unvaccinated individuals but also on consumers who purchased a vaccine but for whom it was ineffective. For $R_0 \geq \rho/(1-\theta)^2$, this second source of externality is still present though the disease is so infectious that all consumers are vaccinated and the first best is obtained.

of taxation, governments would either have to use other instruments, such as bulk purchases, to attain the first best or give up on reaching the first best level of vaccination.

Our results imply that the required subsidy to eradicate a disease is not monotonically increasing in R_0, as might be inferred from the epidemiological literature. The difficulty in addressing a disease does not merely depend on its infectiousness but also on consumers' response to this infectiousness. Consumers respond to extremely infectious diseases by getting vaccinated even if many others also do. But if the disease is only moderately infectious, it is difficult to get large numbers of consumers to buy the vaccine because they depend on others to do so.

4.4 Calibration Illustration

Below we calibrate the model to examine optimal policy in the case of a competitively supplied vaccine or other preventative for which there is an interior solution for use in equilibrium. This is case (iii') in Proposition 4, where $\rho < R_0 \leq \frac{\rho}{1-\theta}$. In this case, in steady-state equilibrium, we can equivalently write

$$MX^* = \frac{\theta h}{\gamma} - c. \tag{18}$$

Thus, under competitive vaccine markets, or more generally for competitively supplied preventative health activities such as water treatment, over this important range of parameter values the marginal externality is equal to the cost of the disease, conditional on infection, times preventative efficacy minus the cost of prevention. Marginal externalities do not vary with R_0 over this range. Under the model, a subsidy of c would be enough to induce consumption of the preventative. Here, c should be interpreted as including any utility cost of prevention.

A quick calibration suggests very large subsidies relative to current levels. To take an example, suppose that circumcision reduces the chance of HIV transmission by $\frac{2}{3}$, and conservatively assume that in a poor country, h for HIV is \$100/year and $\gamma = 0.033$. This implies that the marginal externality from circumcision is $\$2000 - c$. Note that this is a very crude calibration, in the context of a stylized model with homogeneous consumers; as discussed in Section 5, heterogeneity would likely be very important in practice. What this example suggests is that if it cost \$600 to induce someone to be circumcised, that would be "worth it" from a social perspective.

If one assumed an interior equilibrium with homogeneous consumers, and that the lifetime chance of infection in the absence of circumcision were 30%, then the implied c from setting the

marginal private benefit of circumcision equal to its cost might be $\frac{2}{3} \cdot 30\% \cdot \frac{\$100}{0.033} \approx \$600$. Under the model, it would be possible to get to the first best by offering to pay around $600 for circumcision for up to a fraction of each cohort equal to the current proportion immunized plus $\frac{I^*}{.67}$.[13]

Of course, the model we use is extremely stylized, and additional work would be needed to determine if the results generalize to more realistic settings with heterogeneity in transmissibility or heterogeneity in costs of prevention among the population, more realistic birth and death processes, and more complicated disease dynamics such as recovery from disease.

5. Consumer Heterogeneity

Appendix C suggests many results—including our characterization of the equilibrium marginal externality—are robust to incorporating at least some types of heterogeneity, namely heterogeneity in the valuation of health h (for example, due to differences in income). However, the analysis in this section suggests that if there is substantial heterogeneity in c or transmissibility within the population, targeting subsidies will be important.

Consider first heterogeneity in c. Heterogeneity in costs of vaccination, combined with high administrative costs or deadweight loss could weaken the case for subsidies, but it seems unlikely to overturn it. Note that such heterogeneity is plausible. The average cost for polio vaccination, for instance, is probably low, but for those remaining unvaccinated, it may be high. Paying everyone who gets vaccinated the marginal c would be expensive, perhaps explaining the reluctance of some authorities to widely endorse this approach. However, calibration suggests that subsidies would be warranted, if they—such as those in the "lentils for vaccine program" (Banerjee et al. 2010)—could be targeted, even very imprecisely, to people or areas with low vaccine take-up.

Analyzing such cases requires computing the marginal cost and externality benefit of additional vaccinations, taking into account the administrative costs and deadweight loss associated with providing subsidies to infra-marginal consumers. Consider a "lentils for vaccination" program that provided 10 dollars worth of lentils per child immunized (much more than the amount provided in India). Suppose that for every marginal child immunized, nine infra-marginal children received the subsidy, and assume this is considered a pure social cost rather than a transfer[14]—so that the subsidy costs 100 dollars per additional immunization. Even in this case, the program will still be

[13] Under the model, this would reduce prevalence to zero, eliminating the externality benefit of additional circumcisions.

[14] Perhaps a dollar in administrative cost and in deadweight loss is incurred for every dollar of lentils.

justified, as long as vaccine efficacy times the cost of actually getting the diseases prevented by the immunization exceeds 100 dollars plus the one or two dollar marginal manufacturing and delivery costs of the vaccine plus any utility costs of vaccination. This condition seems almost surely met. A similar exercise could be conducted to determine the acceptable rate of non-use in a program that subsidized mosquito nets or chlorine for water purification.

Heterogeneity in infection risk also points to the importance of targeting, but it suggests that a naïve estimate of the marginal externality associated with vaccination that does not take into account population heterogeneity may be a more accurate measure of the marginal externality among high-risk groups than among the population as a whole.

Consider a population consisting of n completely distinct, epidemiologically separate subpopulations with different R_0 and different Q. Consider first the case in which some sub-populations have $R_0 < 1$ and others have $\rho < R_0 < \frac{\rho}{1-\theta}$. In such populations, overall prevalence will be positive, but within sub-populations with $1 - \frac{1}{1-R_0} - \frac{\theta Q}{\gamma} < 0$, there is no social benefit of vaccination. Our model would yield the correct marginal externality of vaccination in the subgroups with $1 - \frac{1}{1-R_0} - \frac{\theta Q}{\gamma} > 0$. However, calculating R_0 based on average prevalence in the society would lead to potentially serious overestimation of the average marginal externality created by immunizing randomly selected susceptibles in the population. Suppose, for example, that the population consists of multiple distinct groups, and that in 90% of the population $1 - \frac{1}{1-R_0} - \frac{\theta Q}{\gamma} < 0$ but that in 10% of the population $1 - \frac{1}{1-R_0} - \frac{\theta Q}{\gamma} > 0$. Based on overall prevalence, a naïve analyst might estimate the marginal externality to be $\theta \frac{h}{\gamma} - c$ in the whole population, when the average marginal externality (*AME*) would be

$$AME = \left(\theta \frac{h}{\gamma} - c\right) \sum_{i=1}^{n} \alpha_i \mathbb{1}\left(1 - \frac{1}{R_0} - \frac{\theta Q_i}{\gamma} > 0\right)$$

where α_i is the population share of subgroup i and $\mathbb{1}(\cdot)$ is an indicator variable, so that the summation equals 0.1 for the case we examine. For these parameter values the *AME* must be less than or equal to the marginal externality associated with the average prevalence in the population. However, the model correctly estimates the marginal externality from immunizations targeted to groups with $I_i = 1$.

We considered above the case in which some subgroups are in case (i) and some are in case (iii), since we think that this will often be the most relevant case. However, theoretically, heterogeneity could also lead to underestimation of the *AME*, if some subgroups were in case (ii) and some were in case (iii). This might be relevant in the case of circumcision, a costly preventative measure, but

is less likely to be the case for most vaccines. If $\theta < 1$ it is also possible that heterogeneity is such that some subgroups are in case (iv), so that the estimated *AME* is greater than the actual marginal externality in any subgroup. In this case, a naive calibration could overestimate the impact of even a well-targeted program. However, this case will be rare if θ is close to one, since case (iv) will arise only for a small set of parameter values.

If there is some transmission across groups and if the disease survives in steady-state in the highest risk group, we conjecture that for realistic parameter values, the benefits of targeting this group will be even greater than in the analysis above, which assumed complete separation.

The average marginal externality created by a subsidy program will depend on whether the program effectively targets those with high risk. If consumers are otherwise homogeneous, differing only in transmission risk, and if they know their transmission risk, high-risk groups will select first into prevention programs. Thus, for example, in African communities with low male circumcision rates, the marginal person induced to use the prevention technology by a subsidy would be a high risk person who would create a large positive externality for others. However, if there were heterogeneity in both transmission risk and cost of vaccination, then much would depend on transmission risk among those who were induced to consume by the subsidy. If low-risk people were also those with a lower cost of prevention, then more subsidies would have to be paid to low-risk people in order to attract the high-risk into prevention.

6. Vaccines Versus Drugs

The previous analysis suggests that governments can use a targeted per-dose subsidy under competition, but that reaching the first-best level of vaccination with per-dose subsidies would be prohibitively expensive under monopoly with direct sales to consumers, and that bulk purchases are a more attractive government policy. Such purchases are in fact common. The huge gap between marginal externalities and the minimum required subsidy creates scope for bargaining. It is natural to assume the threat point in the beginning is a situation of private sales to consumers. We show below that if this is the case, incentives on R&D for vaccines will be lower for an equivalent drug that creates no positive epidemiological externality, even though the social value of the vaccine is greater. This may help explain the general feeling among commentators on the pharmaceutical industry that firms are biased in favor of developing drugs rather than vaccines. By preventing individuals from becoming infected, vaccines interfere with the transmission of the disease, a pos-

itive externality which the firm cannot appropriate. This characteristic distinguishes vaccines from some drugs, which often treat symptoms of the disease without curing it or reducing its transmission. Firms do not simply ignore this externality. Since it lowers other consumers' willingness to pay for their products, it actually reduces their profits. Thus, from an ex ante perspective, firms would have more incentive to invest in developing this type of drug relative to a similarly effective vaccine.

To quantify the monopoly manufacturer's bias toward a drug and against a vaccine, consider a drug that is similar in all ways to the vaccine we have been analyzing except that the drug does not interfere with disease transmission. In particular, suppose both involve the same production cost and efficacy. We will compare the monopolist's profit from the drug and the vaccine, and use the difference as a gauge of the bias manufacturers would have against vaccines.[15]

Finding the right normalization to make drug and vaccine costs equivalent is somewhat delicate because at equal marginal costs of production (*e.g.*, c), the total cost of serving a population with a drug is lower if it only needs to be administered to infected consumers rather than be administered to the whole population ex ante, as is the case with a vaccine. We will finesse this normalization issue by assuming both the drug and vaccine are costless to produce and administer (i.e., $c = 0$). Regarding efficacy, assume that, like the vaccine, the drug treatment is effective with probability θ. Assume that if the drug treatment is effective, it eliminates any harm from the symptoms experienced by infected individuals but does not prevent them from transmitting the disease to susceptible individuals. The consumer only needs one dose of the drug to eliminate symptoms for the rest of her life. If this first dose is ineffective, further doses will be ineffective for her as well.

Computing the steady-state equilibrium in the drug market is straightforward. Note first that if $R_0 \leq 1$, then the disease dies out in the steady-state even without a vaccine, so steady-state equilibrium drug quantity and profit equal zero. Hence, assume $R_0 > 1$. The monopolist can sell the drug for price $P^{d*} = h\theta/\gamma$, the product of the avoided harm h each instant times the remaining lifespan $1/\gamma$ times the probability θ that the drug is effective for the person. The drug can be sold to newly infected consumers each instant. The infection prevalence with a drug, I^d, can be found by substituting $Q = 0$ into equation (5), yielding $I^d = 1 - 1/R_0$. Since I^d is unchanging in the steady-state, the removal of γI^d people from this population due to death by natural causes must be offset

[15] If one takes our focus on the steady-states as literally implying that the discount rate is zero, then there would not be any bias in development decisions. All products producing a positive flow profit would be developed regardless of how large is the up-front development cost and how small is the per-period flow profit; the flow always swamps the up-front cost if the discount rate is zero. We are taking the steady-state profit differential as a rough measure of the present discounted value of profit streams with a positive discount rate.

by an equal number γI^d of new infections each instant. So drug quantity is $Q^{d*} = \gamma(1-1/R_0)$. Since production is costless, drug profit is

$$\Pi^{d*} = P^{d*}Q^{d*} = \theta h\left(1-\frac{1}{R_0}\right). \tag{19}$$

Letting W^{d*} denote flow social welfare in the steady-state equilibrium with a drug,

$$W^{d*} = h(1-I^{d*}+\theta I^{d*}) = h\left(\theta-\frac{\theta}{R_0}+\frac{1}{R_0}\right). \tag{20}$$

Comparing the expressions for drug profit Π^{d*} in equation (19) and to those for vaccine profit Π^* in Proposition 1 leads immediately to the next proposition.[16]

Proposition 6. Suppose that production is costless and that $R_0 > 1$. In the first best, social welfare is weakly greater with a vaccine. At monopoly prices and quantities, the ratio between drug and vaccine profit, $DV^* = \Pi^{d*}/\Pi^*$, is

$$DV^* = \begin{cases} \theta\left(\dfrac{\sqrt{R_0}+1}{\sqrt{R_0}-1}\right) & \text{if } R_0 < (1-\theta)^{-2} \\ \dfrac{(1-\theta)(R_0-1)}{(1-\theta)R_0-1} & \text{if } R_0 \geq (1-\theta)^{-2}, \end{cases} \tag{21}$$

an expression which exceeds 1, is increasing in θ, is independent of γ, and is decreasing in R_0, with $\lim_{R_0\downarrow 1} DV^* = \infty$ and $\lim_{R_0\uparrow\infty} DV^* = 1$.

At monopoly prices and quantities, steady-state equilibrium social welfare is strictly higher with a vaccine than a drug if and only if $R_0 > \theta^2/(1-\theta)^2$ and strictly lower if and only if $R_0 < \theta^2/(1-\theta)^2$.

Figure 4 shows the comparative-static effect of R_0 on DV^*, namely that DV^* is decreasing for all $R_0 > 1$. The horizontal axis is restricted to the values $R_0 > 1$ because DV^* is undefined for $R_0 \leq 1$, being the ratio of 0 over 0. The figure also demarcates the regions in which equilibrium welfare at market prices and quantities from one product exceeds that from the other. Proposition 6 states that the firm strictly prefers to develop the drug rather than the similarly effective vaccine as long as there is a nontrivial market for the products. However, for the range of parameters ($R_0 > \theta^2/(1-\theta)^2$), social welfare is higher with the vaccine. Therefore, for sufficiently high R_0 or sufficiently low θ, the firm develops the "wrong" product. While a first-best quantity of the drug is sold, a given drug dose is socially inferior to a vaccine dose because the drug dose offers no positive externality for other consumers. Thus, our discussion of externalities in the previous

[16]The proof in Appendix A fills in some remaining details.

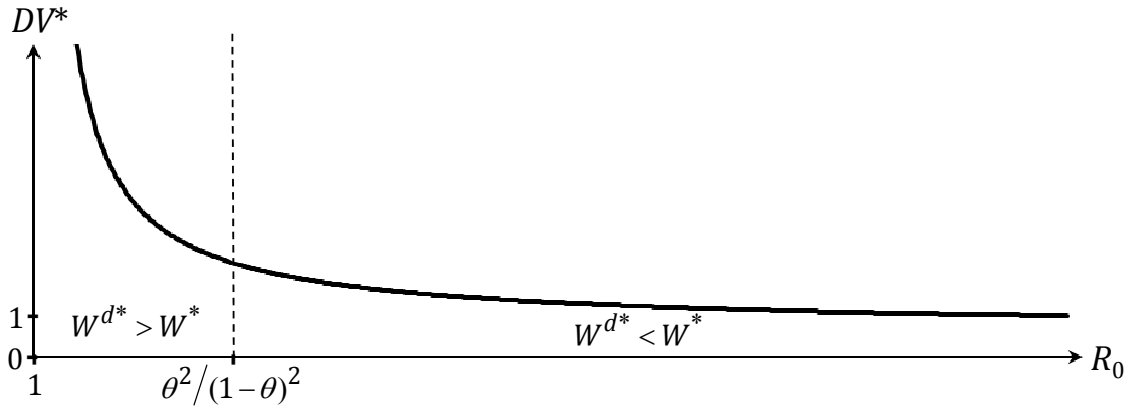

Figure 4: Ratio of drug to vaccine profit, DV^*

two sections provides one rationale for the commonly perceived bias of firms against developing vaccines and toward drugs. Proposition 6 quantifies the size of the bias and characterizes when it should be expected to be largest.

7. Conclusion

We create and calibrate an integrated economic and epidemiological model of disease externalities. We estimate large marginal externalities from disease prevention. Monopoly is particularly costly for vaccines because it creates incentives to restrict supply even if consumers are homogeneous. These distortions are greatest for intermediate disease transmissibility. Consequently, increasing disease transmissibility can reduce prevalence at high levels of transmissibility under monopoly. Under competition, this occurs at lower levels of transmissibility.

Per-dose subsidies are a prohibitively expensive way of reaching optimal vaccination levels under monopoly, but bulk purchases are more attractive. If bargaining over prices for these purchases reflects the threat point of sales to individuals, R&D incentives for vaccines will be lower than for drugs, especially for rare diseases.

Our epidemiological model—although standard—has limitations, lacking heterogeneity across consumers and focusing on the case of a constant hazard rate of death for a non-fatal disease. The virtue of this simplicity is our ability to derive closed form solutions, although we also present an extension to the model that suggests our main results are robust to incorporating at least some types

of heterogeneity. A pair of companion papers (Kremer and Snyder 2015; Kremer and Snyder 2018) examine heterogeneity in more detail, although without integrating epidemiological externalities. Combining heterogeneity and epidemiological externalities is an important topic for future work. It is worth noting that our model also suggests that improvements in vaccine efficacy generate private gains for vaccine manufacturers that are smaller than the social gains. This may provide a rationale for regulatory requirements on minimum vaccine efficacy, which we are currently exploring in related work.

We also focus only on steady-states. Analysis of transition paths awaits future work, but convergence to the steady-state is fairly rapid in many epidemiological steady-state models, so we are optimistic that the key messages will carry over, at least qualitatively.

Another implication of the differences in our results across monopoly and perfectly competitive market structures is that the optimal patent length for vaccines may be shorter than for otherwise equivalent drug treatments. Positive discounting will reduce the optimal price for monopolists, as would a time limit on their monopoly. Having shorter, finite patent lengths for vaccines gives monopolist firms incentives to lower their prices.

Previous work (Kremer and Snyder 2015; Kremer and Snyder 2018) has argued that if infection risk is heterogeneous among the population—and particularly if it is highly skewed—then differences in the timing of the administration of drug treatments and vaccines allow drug manufacturers to extract more rent from consumers than vaccine manufacturers, thus driving a wedge between private and social incentives to invest in vaccine research and development (R&D). That work suggested distortions might be particularly large for HIV.

Both this previous work and the current paper discuss market distortions which suggest a case for subsidies for vaccine R&D. Because HIV is relatively rare, this paper also suggests a strong case for subsidies for activities that reduce HIV transmission in particular. Further analysis would be needed to examine the case of HIV in more detail, since our model is one with a homogeneous population and the low overall prevalence of HIV might be consistent with heterogeneity in risk in which R_0 is large in some populations and small in others. Additional work would be needed to examine the implications of these externalities in a more complicated model with a heterogeneous population.

Appendix A. Proofs

This appendix supplies proofs not included in the text.

Proof of Proposition 1: We here present proofs for the monopoly case in detail. While proofs for the perfect competition case can be proved directly, the direct proofs are omitted for space considerations. Instead, our approach is to note that the formulas are a special case of the Cournot formulas presented in Appendix B, in the limit as $n \to \infty$.

Consider first the case when disease transmissibility is low enough that $R_0 \leq 1$. Then epidemiological relation (5) implies $I^* = 0$ for all $Q \in [0, \gamma]$. Substituting $I = 0$ in (12) and (8) implies $Q^* = 0$ and $MPB^* = 0$. For any positive price, no vaccine is sold, consistent with equilibrium. This establishes case (i).

Next, consider the case in which $1 < R_0 \leq \rho$. To see that $Q^* = 0$, suppose for the sake of contradiction that $Q^* > 0$. Now

$$\hat{I}(P^*) \geq \frac{\gamma c}{(\theta h - \gamma c) R_0} \tag{A1}$$

$$\geq \frac{\gamma c}{\theta h}. \tag{A2}$$

Equation (A1) follows from substituting $c \leq P^*$ into (11). Equation (A2) follows from the upper bound on R_0. Further

$$I^* = \max\left\{0, 1 - \frac{\theta Q^*}{\gamma} - \frac{1}{R_0}\right\} \tag{A3}$$

$$= 1 - \frac{\theta Q^*}{\gamma} - \frac{1}{R_0} \tag{A4}$$

$$\leq 1 - \frac{1}{R_0} \tag{A5}$$

$$\leq \frac{\gamma c}{\theta h}. \tag{A6}$$

Equation (A3) follows from (5). To see (A4) requires a few steps. Note that $I^* \geq \hat{I}(P^*)$ or else (12) would imply $Q^* = 0$. But (A2) implies $\hat{I}(P^*) > 0$. Hence $I^* > 0$, and so I^* must equal the second argument inside the max operator in (A3). Equation (A5) follows from $Q^* > 0$ and (A2) from $R_0 \leq \rho$. But (A2) and (A6) imply $I^* < \hat{I}(P^*)$, in turn implying $Q^* = 0$ by (12). This contradiction proves $Q^* = 0$. Substituting $Q^* = 0$ in (5) gives I^*, and substituting I^* in (8) gives MPB^*. At any price $P^* \geq c$, no vaccine is sold, consistent with equilibrium. This establishes case (ii).

Consider the case in which $R_0 > \rho$ throughout the rest of the proof. As will be seen $Q^* > 0$ in these remaining cases. To solve for Q^*, it is convenient to express the monopolist maximization problem so that the choice variable is quantity rather than price. Inverting (13), we obtain inverse demand

$$P(Q) = \frac{\theta h}{\gamma}\left[1 - \frac{\gamma}{(\gamma - \theta Q) R_0}\right]. \tag{A7}$$

The monopolist chooses Q to maximize $[P(Q) - c]Q$ subject to $Q \leq \gamma$. Applying the Kuhn-Tucker method yields the following solution. If $R_0 < \rho/(1-\theta)^2$, then the constraint does not bind. We

have

$$Q^* = \frac{\gamma}{\theta}\left(1-\sqrt{\frac{\rho}{R_0}}\right), \tag{A8}$$

which upon substitution into (A7) yields

$$P^* = \frac{\theta h}{\gamma}\left(1-\frac{1}{\sqrt{\rho R_0}}\right). \tag{A9}$$

Since consumers are indifferent between buying and not in equilibrium, $MPB^* = P^*$. We can compute Π^* by substituting P^* into profit $[P(Q)-c]Q$. This completes the analysis of case (iii).

If $R_0 \geq \rho/(1-\theta)^2$, then the constraint holds with equality, implying $Q^* = \gamma$. Note that this case can only arise when the vaccine is imperfectly protective so $\theta < 1$. Substituting Q^* into (5) gives I^*, and substituting I^* in (8) gives MPB^*. The monopolist optimally sets $P^* = MPB^*$. We can compute Π^* by substituting P^* and Q^* into profit $[P(Q)-c]Q$. This completes the analysis of case (iv). *Q.E.D.*

Proof of Proposition 2: We here present proofs for the monopoly case in detail; proofs for the perfect competition case are similar.

To show Q^*, MPB^*, and Π^* are nondecreasing in R_0, one can first show that the variables are continuous at the boundaries between all cases (i)–(iv) and then verify that the variables are nondecreasing for each separate case. Form a price function by taking the determinate prices in cases (iii)–(v) and extending this function by taking $P^* = c$ in cases (i) and (ii) in which the equilibrium monopoly price is indeterminate. Similar to the preceding equilibrium variables, one can show that the price function formed in this way is nondecreasing in R_0.

Differentiating the formula for I^* in case (iii) yields

$$\frac{\partial I^*}{\partial R_0} = \frac{1}{R_0^2}\left(1-\frac{\sqrt{\rho R_0}}{2}\right). \tag{A10}$$

Equation (A10) is nonnegative if and only if $R_0 < 4/\rho$ and hence is nonnegative for all R_0 in case (iii) if and only if $2(1-\theta) \geq \rho$. Since I^* is nondecreasing outside of case (iii) and is continuous at the boundaries between cases (i)–(iv), we have that I^* is nondecreasing in R_0 if and only if $2(1-\theta) \geq \rho$. Otherwise, I^* is hump-shaped in R_0 in case (iii), reaching a local maximum of $I^* = \rho/4$ at $R_0 = 4/\rho$. This local maximum is a global maximum if it exceeds all values of I^* in case (iv). Inspection of the stated formula shows that I^* is increasing in R_0 in case (iv), with $\lim_{R_0\to\infty} I^* = 1-\theta$. Hence the local maximum is a global maximum if and only if $\rho \geq 4(1-\theta)$. *Q.E.D.*

Proof of Lemma 1: Epidemiological relation (5) implies

$$HV^* = \begin{cases} h & R_0 \leq 1 \text{ or } I^* = 0 \\ h\left(\dfrac{\theta Q^*}{\gamma} + \dfrac{1}{R_0}\right) & \text{else.} \end{cases} \tag{A11}$$

Assume $R_0 > 1$, $Q^* > 0$, and $P^* > 0$. Then $MPB^* \geq P^* > 0$, implying $I^* > 0$ by (8). Hence, we can eliminate the redundant condition $I^* = 0$ in (A11), yielding equation (16) in the statement of the lemma. *Q.E.D.*

Proof of Proposition 3: The expressions for W^* in the proposition follow immediately from using Lemma 1 to substitute for HV^* and using Proposition 1 to substitute for Q^* in the formula $W^* = HV^* - cQ^*$.

To determine how W^* varies with R_0, one can first verify that W^* is continuous at all the boundaries between cases. One can then focus on how W^* varies with R_0 in each case. It is obvious from the formulas for W^* in the statement of the proposition that W^* is constant in case (i) and decreasing in cases (ii) and (iv). It remains to determine the behavior of W^* in case (iii). Differentiating the formula for W^* in case (iii) yields

$$\frac{\partial W^*}{\partial R_0} = \frac{h}{2R_0^2\sqrt{\rho}}\left(\sqrt{R_0} - 2\sqrt{\rho}\right). \tag{A12}$$

This derivative is nonpositive for all R_0 in case (iii) if $\theta \leq 1/2$. Otherwise, W^* reaches a local minimum of $W^* = 3h/4\rho$ at $R_0 = 4\rho$. This local minimum is a global minimum if it less than all values of W^* in case (iv). Since W^* is decreasing in case (iv), we examine its limiting behavior: $\lim_{R_0 \to \infty} W^* = \theta h/\rho$. Hence the local maximum is a global maximum if and only if $3h/4\rho \leq \theta h/\rho$, or $\theta \geq 3/4$. *Q.E.D.*

Proof of Proposition 4: The text explains the derivation of MX^* in Proposition 4. It remains to use Proposition 4 to characterize the comparative-statics effect of changes in R_0 on MX^*. It can be verified that $MX^* = 0$ for R_0 in case (i), MX^* jumps upward at $R_0 = 1$, and is nonincreasing in R_0 for $R_0 > 1$. This last point can be seen by first showing that MX^* is continuous in R_0 at the boundaries between cases (ii)–(iv) and then showing for each of the cases (ii)–(iv) that MX^* is nonincreasing in R_0. Finally, one can verify that MX^* is strictly decreasing in R_0 in case (ii), establishing that MX^* reaches a unique global supremum in R_0 in the limit $R_0 \downarrow 1$. *Q.E.D.*

Proof of Proposition 5: The first best is achieved without any subsidy in case (i)—because the disease disappears in the steady-state without a vaccine—and case (iv)—because all consumers are vaccinated in equilibrium without a subsidy. It remains to characterize the minimum optimal subsidy for R_0 in cases (ii) and (iii). For the remainder of the proof, assume R_0 falls into one of these cases; i.e., assume $R_0 \in (1, \rho/(1-\theta)^2]$.

We begin by showing that GS^* either results in the disease being eradicated or, failing that, results in all individuals being vaccinated. If $I > 0$, then $MHV = \theta h/\gamma > c$, where the first equality holds by differentiating (A11) with respect to Q and noting $R_0 > 1$ and the second inequality holds by assumption (10). Since the marginal health value exceeds the marginal social cost, the subsidy should be increased until

$$Q = \frac{\gamma}{\theta}\left(1 - \frac{1}{R_0}\right), \tag{A13}$$

the value of Q for which $I = 0$ if it is feasible to increase Q this high. If $R_0 \leq 1/(1-\theta)$, then (A13) does not exceed γ, so it is indeed feasible to increase Q to (A13) and thus to reduce I to 0. If

$R_0 > 1/(1\theta)$, then GS^* induces as high a Q as possible, i.e., $Q = \gamma$.

The proof concludes by calculating the minimum subsidy needed to accomplish the goals set out in the previous paragraph. In the presence of the subsidy, the monopolist maximizes $[P(Q) - \tilde{c}]Q$, which is the same problem as in the proof of Proposition 1 except $\tilde{c} = c - GS$ replaces c. Similar calculations can be used to derive the interior solution for monopoly quantity

$$Q = \frac{\gamma}{\theta}\left[1 - \sqrt{\frac{\theta h}{(\theta h - \gamma\tilde{c})R_0}}\right]. \tag{A14}$$

The interior solution is relevant because the *minimum* optimal subsidy would never be increased beyond the point where the constraint $Q \leq \gamma$ starts to have a positive shadow price. Setting (A14) equal to (A13), substituting $\tilde{c} = c - GS$, and solving for GS yields the expression for GS^* in the statement of the proposition for $R_0 \leq 1/(1-\theta)$. Setting (A14) equal to γ, substituting $\tilde{c} = c - GS$, and solving for GS yields the expression for GS^* in the statement of the proposition for $R_0 > 1/(1-\theta)$. *Q.E.D.*

Proof of Proposition 6: We will verify the last statement of the proposition, which compares social welfare from a vaccine and a drug. The other statements in the proposition follow fairly immediately from the text before the proposition and the statement of the proposition itself. Assume $R_0 > 1$ throughout the proof.

We can compute $W^{d*} - W^*$ using the formula for W^* from Proposition 1 and for W^{d*} from (20). There are two cases to consider.For $1 < R_0 \leq (1-\theta)^{-2}$,

$$W^{d*} - W^* = h\left(\theta - \frac{\theta}{R_0} - 1 + \frac{1}{\sqrt{R_0}}\right). \tag{A15}$$

For $R_0 > (1-\theta)^{-2}$,

$$W^{d*} - W^* = \frac{-\theta h}{R_0}. \tag{A16}$$

It is obvious that $W^* > W^{d*}$ in (A16). One can show $W^{d*} > W^*$ in (A16) if $R_0 < \theta^2/(1-\theta)^2$ and $W^* > W^{d*}$ if the reverse inequality holds strictly. Putting the cases together gives the desired result. *Q.E.D.*

Appendix B. Cournot Competition

This appendix provides a complete analysis of the case of Cournot competition.

Assume the vaccine is manufactured by $n = 1$ homogeneous Cournot firms, who choose quantities each period simultaneously. The analysis is similar to that in Section 1. In particular, cases (i) and (ii) from Proposition 1 will be identical under Cournot. We will focus on the remaining cases in the rest of this appendix.

The inverse demand function is the same as under monopoly, given by (A7). Firm i's profit equals $[P(q_i + Q_{-i}) - c]q_i$, where q_i is firm i's output and Q_{-i} is the output of i's rivals. Taking the first-order condition with respect to q_i and then imposing symmetry by substituting $q_i^* = Q^*/n$ and

$Q^*_{-i} = (n-1)Q^*/n$ yields equilibrium market output

$$Q^* = \frac{\gamma}{\theta}\left(1-\psi\sqrt{\frac{\rho}{R_0}}\right), \tag{A17}$$

where

$$\psi = \frac{n-1}{2n}\sqrt{\frac{\rho}{R_0}}+\sqrt{\left(\frac{n-1}{2n}\right)^2\frac{\rho}{R_0}+\frac{1}{n}}. \tag{A18}$$

Substituting Q^* into equation (5) yields

$$I^* = \psi\sqrt{\frac{\rho}{R_0}}-\frac{1}{R_0}. \tag{A19}$$

Substituting Q^* into (A7) yields

$$P^* = MPB^* = \frac{\theta h}{\gamma}\left(1-\frac{1}{\psi\sqrt{\rho R_0}}\right). \tag{A20}$$

Substituting P^* and Q^* into the profit equation $\Pi = (P^*-c)Q^*$ and rearranging yields equilibrium industry profit

$$\Pi^* = \frac{h}{\rho}\left(1-\frac{1}{\psi}\sqrt{\frac{\rho}{R_0}}\right)\left(1-\psi\sqrt{\frac{\rho}{R_0}}\right). \tag{A21}$$

Substituting (16) and Q^* into $W^* = HV^*-cQ^*$ yields

$$W^* = \frac{h}{\rho}\left(1-\psi\sqrt{\frac{\rho}{R_0}}+\frac{\rho}{R_0}\right). \tag{A22}$$

Finally, substituting (8) and (A20) into $MX^* = MHV^*-MPB^*$ yields

$$MX^* = \frac{\theta h}{\psi\gamma\sqrt{\rho R_0}}. \tag{A23}$$

The preceding analysis holds if $R_0 \leq \psi^2\rho/(1-\theta)^2$. If $R_0 > \psi^2\rho/(1-\theta)^2$, we have $Q^* > \gamma$ for the Q^* in (A17). Producing more than the number of consumers would result in a market price of zero and zero profits for all firms. Instead, firms produce an equal share of industry output $Q^* = \gamma$. The rest of the equilibrium variables have the same formula as in case (iv) of Proposition 1. Note that the interval in which the case applies is different.

It is easily seen that $\psi = 1$ for $n = 1$, and thus that the preceding expressions for the equilibrium variables collapse to their monopoly values given in Proposition 1. It is also easily seen that $\psi = \sqrt{\rho/R_0}$ in the limit as $n \to \infty$, and thus that the preceding expressions for the equilibrium variables collapse to their values under perfect competition given in Proposition 1.

Appendix C. Heterogeneous Consumers

This appendix shows that our central results extend to the case among consumers with heterogeneous valuation of health, for example, due to difference in income. In particular, we will focus on the result that the marginal externality MX^* reaches a maximum at an interior value of R_0.

Assume consumer i's health benefit is a continuous random variable h_i with distribution function F, density f, and full support $[0,\infty)$. The assumption of full support ensures that, for any given price, there will always be some consumers with a willingness to pay above this price and some with a willingness to pay below this price, ruling out corner solutions. Assume indifferent consumers do not purchase the vaccine; since such consumers will have measure 0 this is not an important assumption but allows us to make clean "if and only if" statements. Denote the complement of F by $\bar{F}(h_i) = 1 - F(h_i)$.

Following equation (8), consumer i's marginal private benefit is

$$MPB_i = \frac{\theta h_i I R_0}{\gamma(1+IR_0)}. \tag{A24}$$

She purchases the vaccine if and only if $MPB_i > P$ or equivalently if and only if $h_i > \hat{h}(I,P)$, where

$$\hat{h}(I,P) = \frac{\gamma P(1+IR_0)}{\theta IR_0}. \tag{A25}$$

The structural demand relation is given by

$$Q = \gamma \bar{F}(\hat{h}(I,P)), \tag{A26}$$

the product of the flow of new consumers, γ, and the probability an individual consumer purchases $\bar{F}(\hat{h}(I,P))$. Substituting for I from (5) into (A26) we obtain[17]

$$Q = \gamma \bar{F}\left(\frac{\gamma P R_0(\gamma - \theta Q)}{\theta[R_0(\gamma - \theta Q) - \gamma]}\right). \tag{A27}$$

Reduced-form demand $D(P)$ is given by the implicit solution for Q in (A27). The monopoly's equilibrium price P^* maximizes profit $(P-c)D(p)$. Rather than working through the derivation of all steady-state-equilibrium variables, we will focus the analysis on characterizing the marginal externality MX^*. With homogeneous consumers, welfare and externality calculations depended only on the number of vaccinated/unvaccinated consumers. With heterogeneous consumers, this is no longer the case because vaccinated consumers form a self-selected sample, having high values of h_i. Welfare calculations are complicated by the need to keep track of the sample of consumers. The next proposition provides an expression for MX^* under consumer heterogeneity and characterizes the comparative statics of MX^* with respect to R_0.

Proposition 7. Suppose consumers have heterogeneous health benefits h_i. The marginal external-

[17] We can ignore the $I = 0$ branch in (5) because we will focus on the $R_0 > 1$ case in the analysis. The $R_0 \leq 1$ case is the same as in the homogeneous-consumer case, with the disease dying out in the steady-state.

ity in steady-state equilibrium is

$$MX^* = \gamma \left[\frac{\theta R_0}{\gamma^2 (1+IR_0)^2} \right] \left[E(h_i) - \theta \int_{\hat{h}(I^*,P^*)}^{\infty} h_i f(h_i)\, dh_i \right], \tag{A28}$$

which has an interior maximum in R_0.

The expression for MX^* in (A28) is the product of three factors. The first, γ, accounts for the mass of new consumers each instant. The next factor captures the effect of an additional vaccination on the length of time that other unvaccinated people remain healthy on average. The last factor scales the health duration by the health benefit, accounting for the selection effect that vaccinated consumers are the highest benefit ones.

The fact that MX^* reaches an interior maximum in R_0 is established by showing that $MX^* = 0$ for $R_0 \leq 1$, then MX^* jumps to $\theta E(h_i)/\gamma$ at $R_0 = 1$. While we cannot determine what MX^* looks like without a specific functional form for the distribution of types F, for any F we know that MX^* eventually asymptotes to 0 for sufficiently large R_0.

Proof of Proposition 7: If $R_0 \leq 1$, then the disease dies out even without a vaccine, implying $MX^* = 0$. So suppose $R_0 > 1$ for the rest of the proof.

We first compute the health value (in the sense of the social benefit) HV^* in the heterogeneous-consumer case by computing the expected health benefit over the lifespan of a single consumer of type h_i and then integrating over possible types to arrive at the lifetime health benefit for a cohort of newborns born at time t. This is the same as the health value HV^* for the population at instant t.

The expected health benefit for an unvaccinated consumer is

$$\int_0^{\infty} h_i e^{-(\gamma+\gamma IR_0)t} dt = \frac{h_i}{\gamma(1+IR_0)}, \tag{A29}$$

On the left-hand side of (A29), h_i is the benefit consumer i receives from being healthy at instant t and $e^{-(\gamma+\gamma IR_0)t}$ is i's probability that i survives to instant t, neither dying of natural causes nor becoming infected before then, conditional on not being vaccinated. The expected health benefit for a vaccinated consumer is

$$\int_0^{\infty} h_i(1+\theta IR_0) e^{-(\gamma+\gamma IR_0)t} dt \tag{A30}$$

$$= \frac{h_i(1+\theta\gamma IR_0)}{\gamma(1+IR_0)}. \tag{A31}$$

Equation (A30) differs from the left-hand side of (A29) only in the added term in parentheses $\theta\gamma IR_0$. This is the probability that the consumer would have been infected in the absence of a vaccine at instant t (probability γIR_0) but was protected by the vaccine (probability θ). Integrated across consumer types and scaled by the mass of newborns each instant, the health benefit for a

cohort of newborns is

$$HV = \gamma \left\{ \int_0^{\hat{h}} \frac{h_i}{\gamma(1+IR_0)} f(h_i)\, dh_i \right. \tag{A32}$$
$$\left. + \int_{\hat{h}}^{\infty} \frac{h_i(1+\theta\gamma IR_0)}{\gamma(1+IR_0)} f(h_i)\, dh_i \right\}$$
$$= \frac{E(h_i)}{1+IR_0} + \frac{\theta IR_0}{1+IR_0} \int_{\hat{h}}^{\infty} h_i f(h_i)\, dh_i, \tag{A33}$$

where $\hat{h}$ is shorthand for $\hat{h}(I,P)$.

To find *MHV*, we need to differentiate (A33) with respect to Q. Some preliminary steps will help compute this derivative. Expressed in terms of Q,

$$\frac{1}{1+IR_0} = \frac{\gamma}{(\gamma-\theta Q)R_0}, \tag{A34}$$

implying

$$\frac{\partial}{\partial Q}\left(\frac{1}{1+IR_0}\right) = \frac{\theta\gamma}{(\gamma-\theta Q)^2 R_0} \tag{A35}$$
$$= \frac{\theta R_0}{\gamma(1+IR_0)^2}. \tag{A36}$$

Similarly,

$$\frac{\partial}{\partial Q}\left(\frac{IR_0}{1+IR_0}\right) = -\frac{\theta R_0}{\gamma(1+IR_0)^2}. \tag{A37}$$

Differentiating (A26) yields $dQ/d\hat{h} = -\gamma f(\hat{h})$, implying

$$\frac{d\hat{h}}{dQ} = -\frac{1}{\gamma f(\hat{h})} \tag{A38}$$

by the Inverse Function Rule. Substituting (A36), (A37), and (A38) into the derivative of (A33) and rearranging yields

$$MHV^* = \gamma \left[\frac{\theta R_0}{\gamma^2(1+IR_0)^2}\right] \times \left[E(h_i) - \theta \int_{\hat{h}(I^*,P^*)}^{\infty} h_i f(h_i)\, dh_i\right] + c. \tag{A39}$$

Combining (A39) with the fact that $MX^* = MHV^* - MPB^* = MHV^* - c$ yields (A28).

To determine the limiting behavior of MX^* as $R_0 \to \infty$, first notice

$$MX^* \leq \frac{\theta E(h_i) R_0}{\gamma(1+IR_0)^2} \tag{A40}$$

$$= \frac{\theta E(h_i)}{\gamma R_0[1-\theta \bar{F}(\hat{h}(I^*,P^*))]^2} \tag{A41}$$

$$< \frac{\theta E(h_i)}{\gamma(1-\theta)^2 R_0}, \tag{A42}$$

where (A40) follows from (A28), (A41) follows from (5) and (A26), and (A42) follows because $\bar{F}(\hat{h}) < 1$. But (A42) implies

$$\lim_{R_0 \uparrow \infty} MX^* \leq \lim_{R_0 \uparrow \infty} \frac{\theta E(h_i)}{\gamma(1-\theta)^2 R_0} = 0.$$

Q.E.D.

References

Ahituv, Avner, V. Joseph Holz, and Tomas Philipson. (1996) "The Responsiveness of the Demand for Condoms to the Local Prevalence of AIDS," *Journal of Human Resources* 31(4): 869–897.

Althouse, Benjamin M., Theodore C. Bergstrom, and Carl T. Bergstrom. (2010) "A Public Choice Framework for Controlling Transmissible and Evolving Diseases,âĂİ *Proceedings of the National Academy of Sciences* 107(1): 1696.

Anderson, Roy and Robert May. (1991) *Infectious Diseases of Humans*. Oxford: Oxford University Press.

Bailey, Norman. (1975) *The Mathematical Theory of Infectious Diseases and Its Application*. London: Griffen.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Dhruva Kothari (2010) "Improving Immunization Coverage in Rural India: A Clustered Randomized Controlled Evaluation of Immunization Campaigns with and without Incentives," *British Medical Journal* 340.

Berndt, Ernst, Rachel Glennerster, Michael Kremer, Jean Lee, Ruth Levine, Georg Weisacker, and Heidi Williams. (2007) "Advance Market Commitments for Vaccines Against Neglected Diseases: Estimating Costs and Effectiveness," *Health Economics* 16: 491–511.

Bethune, Zachary and Anton Korinek (2020) "Covid-19 Infection Externalities: Trading Off Lives vs. Livelihoods," National Bureau of Economic Research Working Paper. (No. w27009).

Boulier, Bryan. (2010) "A Shot in the Dark: Uncertainty and Vaccine Demand and Supply," George Washington University working paper.

Boulier, Bryan, Tejwant Datta, and Robert Goldfarb. (2007) "Vaccination Externalities," *The B.E. Journal of Economic Analysis & Policy* 7: www.bepress.com/bejeap/vol7/iss1/art23.

Brennan, Michael J. (2009) "The US Food and Drug Administration Provides a Pathway for Licensing Vaccines for Global Diseases," *PLoS Medicine* 6(7): e1000095.

Brito, Dagobert L., Eytan Sheshinski, and Michael D. Intrilligator. (1991) "Externalities and Compulsory Vaccination," *Journal of Public Economics* 45: 69–90.

Chen, Frederick and Flavio Toxvaerd. (2014) "The Economics of Vaccination." *Journal of Theoretical Biology* 363: 105–117.

Chowell, Gerardo, Paul W. Fenimore, Melissa A. Castillo-Garsow, and Carlos Castillo-Chavez. (2003) "SARS Outbreaks in Ontario, Hong Kong and Singapore: The Role of Diagnosis and Isolation As a Control Mechanism," *Journal of Theoretical Biology* 224: 1–8.

Cook, Joseph, Marc Jeuland, Brian Maskery, Donald Lauria, Dipika Sur, John Clemens, and Dale Whittington. (2009) "Using Private Demand Studies to Calculate Socially Optimal Caccine Subsidies in Developing Countries.âĂİ *Journal of Policy Analysis and Management* 28 (1): 6–28.

Francis, Peter J. (1997) "Dynamic Epidemiology and the Market for Vaccinations," *Journal of Public Economics* 63: 383–406.

Funk, Sebastian, Marcel Salathé, and Vincent A. A. Jansen. (2010) "Modelling the Influence of Human Behaviour on the Spread of Infectious Diseases: A Review,âĂİ *Journal of The Royal Society Interface* 7 (50): 1247–56.

Galeotti, Andrea, and Brian W. Rogers. (2013) "Strategic Immunization and Group Structure," *American Economic Journal: Microeconomics* 5(2): 1–32.

Geoffard, Pierre-Yves and Tomas Philipson. (1997) "Disease Eradication: Public vs. Private Vaccination," *American Economic Review* 87: 222–230.

Gersovitz, Mark (2003) "Births, Recoveries, Vaccinations, and Externalities," in R. Arnott (ed.) *Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz*, pp. 469–483.

Gersovitz, Mark and Jeffrey S. Hammer. (2004) "The Economical Control of Infectious Diseases," *Economic Journal* 114: 1–27.

Gersovitz, Mark and Jeffrey S. Hammer. (2005) "Tax/Subsidy Policy Toward Vector-Borne Infectious Diseases," *Journal of Public Economics* 89: 647–674.

Kremer, Michael. (1996) "Integrating Behavioral Choice into Epidemiological Models of AIDS," *Quarterly Journal of Economics* 111(2): 549–573.

Kremer, Michael and Christopher M. Snyder. (2015) "Preventives vs. Treatments," *Quarterly Journal of Economics*130(3): 1167–1239

Kremer, Michael and Christopher M. Snyder. (2018) "Preventives versus Treatments Redux: Tighter bounds on Distortions in Innovation Incentives with an Application to the Global Demand for HIV Pharmaceuticals," *Review of Industrial Organization* 53 (1): 235–273

Manfredi, Piero, and Alberto D'Onofrio (ed.) (2013) *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases*. New York: Springer-Verlag.

Mechoulan, Stéphane. (2007) "Market Structure and Communicable Diseases," *Canadian Journal of Economics* 40: 468–492.

Philipson, Tomas. (2000) "Economic Epidemiology and Infectious Disease," in J. Newhouse and T. Culyer (ed.) *Handbook of Health Economics, Volume IB*: 443–465.

Thomas, Patricia. (2002) "The Economics of Vaccines," Harvard Medical International (HMI) World. September/October.

U.S. General Accounting Office. (2002) "Childhood Vaccines: Ensuring an Adequate Supply Poses Continuing Challenges," Report no. GAO-02-987.