International Journal of Industrial Organization 84 (2022) 102840

Contents lists available at ScienceDirect

International Journal of Industrial Organization

journal homepage: www.elsevier.com/locate/ijio

Optimal vaccine subsidies for endemic diseases☆

Matthew Goodkin-Gold[a], Michael Kremer[b], Christopher M. Snyder[c,*], Heidi Williams[d]

[a] Department of Economics, Harvard University, Cambridge, Massachusetts, USA
[b] Department of Economics, University of Chicago, Chicago, Illinois, USA
[c] Department of Economics, Dartmouth College, Hanover, New Hampshire, USA
[d] Department of Economics, Stanford University, Stanford, California, USA

ARTICLE INFO

Article history:
Available online 2 April 2022

JEL classification:
O31
L11
I18
D42

Keywords:
Vaccine
Epidemiology
Externality
Pharmaceutical

ABSTRACT

In Goodkin-Gold et al. (2021), we analyzed optimal subsidies for a vaccine against an epidemic outbreak like Covid-19. This companion paper alters the underlying epidemiological model to suit endemic diseases requiring continuous vaccination of new cohorts—also suiting an epidemic like Covid-19 if, following Gans (2020), one assumes peaks are leveled by social distancing. We obtain qualitatively similar results: across market structures ranging from perfect competition to monopoly, the subsidy needed to induce first-best vaccination coverage on the private market is highest for moderately infectious diseases, which invite the most free riding; extremely infectious diseases drive more consumers to become vaccinated, attenuating externalities. Stylized calibrations to HIV, among other diseases, suggest that first-best subsidies can be exorbitantly high when suppliers have market power, rationalizing alternative policies observed in practice such as bulk purchases negotiated by the government on behalf of the consumers.

1. Introduction

Vaccines exert a positive externality by reducing the spread of a disease from the vaccinated to others. In a companion paper (Goodkin-Gold et al., 2021), we related characteristics of the disease to the size of the externality and the optimal

☆ This paper is one of two companion papers derived from a longer National Bureau of Economic Research working paper, no. 28085, "Optimal Vaccine Subsidies for Endemic and Epidemic Diseases." Snyder drew on this paper and other joint work for his presentation in the invited session on "Policies for Vaccines" at the 2021 EARIE meetings, chaired by Kurt Brekke. The article was substantially improved following insightful suggestions from an anonymous referee and the editor, Elisabetta Iossa. The authors are grateful for helpful comments from Chris Avery, Witold Więcek, and participants in the EARIE invited session, Harvard Economics Department seminar, Yale School of Medicine seminar, Cornell University's "Infectious Diseases in Poor Countries and the Social Sciences" conference, DIMACS "Game Theoretic Approaches to Epidemiology and Ecology" workshop at Rutgers University, "Economics of the Pharmaceutical Industry" roundtable at the Federal Trade Commission's Bureau of Economics, 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 sessions on "Economics of Infectious Disease" and "Private and Social Returns to R&D and Vaccine Development," Health and Pandemics (HELP!) Economics Working Group "Covid-19 and Vaccines" workshop, and NBER "Covid-19 and Health Outcomes" conference. Christopher Cardillo, Nishi Jain, Amrita Misha, Ralph Skinner, and Alfian Tjandra provided excellent research assistance. Williams gratefully acknowledges financial support from NIA grant number T32–AG000186 to the NBER.

* Corresponding author.
E-mail address: chris.snyder@dartmouth.edu (C.M. Snyder).

https://doi.org/10.1016/j.ijindorg.2022.102840

subsidy needed to correct it. We analyzed a model of the vaccine market in which demanders and suppliers base their economic decisions on rational expectations of the infection risk derived from a standard susceptible-infected-removed (SIR) epidemiological model. Within the SIR framework, we made specific modeling choices to suit an intensive vaccine campaign against an epidemic disease such as Covid-19 that rises and falls within years rather than decades.

This paper revisits the analysis of vaccine externalities and optimal subsidies, maintaining the SIR framework, but making specific modeling choices to suit endemic diseases such as HIV or measles remaining in the population over the span of generations. Instead of an intensive vaccine campaign conducted at the outset of the epidemic studied in the companion paper, here new cohorts are continuously vaccinated, leading to a constant infection rate in the steady-state equilibrium.

The epidemiological model adopted here may apply beyond endemic diseases, to epidemic diseases, once endogenous social distancing is considered. Social distancing by fully rational agents is notoriously difficult to formalize because the level of distancing depends not just on the current infection level but the future path of the epidemic, which in turn depends on the path of distancing decisions.[1] For example, agents may not bother to distance regardless of the current infection level if they expect the infection rate to rise to a level that will make infection hard to avoid in any event. Gans (2020) suggests that social distancing will iron out an epidemic's peaks and extend its duration, leading to a constant infection rate over an extended period. The steady state we analyze has a constant infection rate by definition, so can be interpreted as approximating the path of an epidemic disease with endogenous social distancing.

Despite deriving from different epidemiological models, key qualitative results carry over from the companion paper. If anything, the results are clearer here because the steady-state values of equilibrium variables admit closed-form solutions, which can be differentiated to determine comparative statics. We continue to find that, under market structures ranging from perfect competition to monopoly, the marginal externality and optimal subsidy necessary to correct it are hump shaped in the disease's basic reproductive number, $\mathscr{R}_0$, a measure of the disease's infectiousness. No externality arises for diseases with extremely low $\mathscr{R}_0$ since they die out in the steady state without a vaccine. The externality also disappears for diseases with extremely high $\mathscr{R}_0$ since all consumers are driven to purchase the vaccine in that case. An unvaccinated person is almost certain to contract a highly infectious disease even if all others are vaccinated since even a highly effective vaccine cannot prevent all of them from transmitting the disease if it is imperfect.

We find that the vaccine market performs much worse under monopoly than under perfect competition. The distortion due to market power compounds the distortion due to free riding, leading to large deadweight losses for some parameters. It is difficult to overcome the distortions with a subsidy because the monopoly absorbs subsidy increases, passing little on to consumers. For certain parameter limits, the subsidy required to induce the first-best vaccination level can be unboundedly high. Practical constraints on paying exorbitant subsidies may call for other policies such as bulk purchases negotiated by the government on behalf of the consumers. Given the choice of developing a vaccine or a similarly effective drug, a monopoly would be biased toward the product with less free riding—the drug to the extent it treats symptoms without reducing transmission. The monopoly bias toward a drug exhibits the same hump shape in $\mathscr{R}_0$ as the marginal externality and optimal subsidy on the vaccine market. Whether the bias is socially harmful depends on the parameters: the drug market has fewer distortions, but the drug lacks the vaccine's positive externality in terms of reducing transmission.

Under all market structures we study, we find the possibility of an epidemiological version of the Peltzman (1975) effect. The Peltzman effect occurs when an increase in infectiousness $\mathscr{R}_0$ reduces the equilibrium infection rate due to offsetting risk compensation by consumers, who increase their vaccination rate. The Peltzman effect never leads to an increase in equilibrium welfare under perfect competition but can under monopoly because of the compounded distortions in that market structure.

As a capstone to the theoretical analysis, we provide calibrations to the case of HIV. At the calibrated parameter values, a hypothetical vaccine market would generate the first best if supplied by perfect competitors. For a hypothetical monopoly vaccine market, the estimate of $\mathscr{R}_0 = 4$ for HIV from the epidemiological literature is a sort of worst case: the infection rate and deadweight loss are maximized precisely at $\mathscr{R}_0 = 4$, requiring massive subsidies to obtain the first best under monopoly. We also provide calibrations to measles and SARS, more and less infectious diseases than HIV by way of contrast.

Turning to a review of related literature,[2] our most direct contribution is to the theoretical literature on vaccine externalities[3] by economists[4] and mathematical epidemiologists.[5] Our paper is close to Althouse et al. (2010), which also analyzes welfare and optimal subsidies on the vaccine market. We build on their work, allowing for imperfect vaccines and supplier market power and deriving comparative statics from closed-form expressions rather than calibrations. Our paper is perhaps closest to two companion papers, Mamani et al. (2012) and Adida et al. (2013), which also analyze optimal vaccine subsidies for various degrees of supplier market power. However, those papers do not analyze how equilibrium variables vary with $\mathscr{R}_0$. Thus, our key results on the nonmontonicity of optimal subsidies in $\mathscr{R}_0$ and other comparative statics are novel (as our

[1] Recent theoretical advances include Acemoglu et al. (2020), Atkeson et al. (2020, 2021), Farboodi et al. (2020), Jones et al. (2020), Keppo et al. (2021), Makris and Toxvaerd (2020), McAdams (2020), Rachel (2020), Tröger (2020), and Toxvaerd (2020). Early work includes Kremer (1996), Regula (2010), Fenichel (2013), and Toxvaerd (2019).

[2] See our companion paper for an extensive review. Our review here is limited to studies that most directly relate to the contributions of this paper.

[3] See surveys by Philipson (2000), Manfredi and D'Onofrio (2013), Avery et al. (2020), and McAdams (2021).

[4] In addition to the work cited in footnote [1], see Brito et al. (1991), Chen and Toxvaerd (2014), Francis (1997), Geoffard and Philipson (1997), Gersovitz (2003), Gersovitz and Hammer (2004, 2005), Boulier et al. (2007), Manski (2010, 2017, 2021), and Galeotti and Rogers (2013).

[5] See Bauch and Earn (2004), Funk et al. (2010), and Manfredi and D'Onofrio (2013).

results on suppliers' bias against vaccines toward drugs and our calibrations). Our paper also relates to the growing body of recent work spurred by the Covid-19 pandemic to improve the epidemiological models used for economic forecasting including Acemoglu et al. (2021), Alvarez et al. (2021), Avery (2021), Fajgelbaum et al. (2021), Eichenbaum et al. (2020), and Ellison (2020), in addition to work cited in footnote 1.

The remainder of the paper is structured as follows. Section 2 outlines the model. Section 3 characterizes the equilibrium and derives its comparative statics under perfect competition and monopoly. Section 4 analyzes optimal subsidies. Section 5 examines a firm's incentives to develop a vaccine relative to a drug that does not prevent transmission. Section 6 provides calibrations of the model to HIV, measles, and SARS. Section 7 concludes. The more technical material is relegated to a series of appendices. Appendix A provides omitted proofs, Appendix B extends the results to Cournot competition among a general number of firms n, and Appendix C extends the results to heterogeneous consumers.

2. Model

We begin by specifying the epidemiological model and solving for the steady-state infection rate as a function of immunizations. We then introduce agents on the demand and supply side of the vaccine market who based their rational economic decisions on the prevailing infection rate, determining the equilibrium vaccine quantity. When fed back into the epidemiological model, the equilibrium vaccine quantity must generate an infection rate validating agents' rational expectations, closing the model.

2.1. Epidemiology

We use the standard susceptible-infected-removed (SIR) epidemiological model (Kermack and McKendrick, 1927), augmented to reflect demographic flows (births and deaths). Time is continuous, indexed by t. Let $\mu \in (0, 1)$ denote the mortality rate from causes other than the disease, referred to as natural causes. For simplicity, we will study the case of a nonfatal disease and set the birth rate to equal to μ, leaving the population size constant over time. Anticipating their role in the vaccine market, we refer to individuals in this population as consumers.

Consumers are partitioned into four compartments according to their current state: susceptible to infection S_t, infected I_t, recovered from an infection R_t, or successfully immunized Z_t. Normalizing the mass of consumers to 1,

$$S_t + I_t + R_t + Z_t = 1, \tag{1}$$

allows compartments to be interpreted as either masses or proportions.

Turn next to the equations determining the evolution of compartments over continuous time, starting with Z_t. Let $Q \geq 0$ be the quantity of vaccine purchased each instant. For now take Q as given; later, we will solve for its equilibrium value using the economic model and substitute this value back into the epidemiological model. In principle, Q could vary over time, but we omit a time subscript anticipating that we will solve for its equilibrium value in the steady state. Let $\theta \in (0, 1)$ denote vaccine efficacy. Assume that vaccinations are adminstered to newborns[6] and that if the initial dose is not effective for a person, further doses will not be effective for that person either. The mass of newborns who are successfully immunized each instant is θQ. Letting dots denote time derivatives, the rate of change of the successfully immunized population is

$$\dot{Z}_t = \theta Q - \mu Z_t, \tag{2}$$

the number of newborns who are immunized minus the number of immunized individuals who die of natural causes.

The rate of change of the infected population is

$$\dot{I}_t = \beta I_t S_t - (\alpha + \mu) I_t. \tag{3}$$

The first term captures the flow of new infections. A susceptible consumer contracts the disease from an infected consumer at rate $\beta > 0$, embodying the rate of contact between people and the rate at which a contact leads to infection. Assuming the infection rate is linear in the number of infected consumers, a single susceptible consumer is infected with probability βI_t, and the mass of susceptibles generates $\beta I_t S_t$ new infections. The mass of infected consumers is reduced by the mass αI_t of them who recover, where $\alpha \in (0, 1)$ denotes the recovery rate, and by the number of infected consumers who die. Consistent with our assumption that the disease is nonfatal, the mortality rate is the same μ for infected consumers as for other compartments.

Assuming that recovered people cannot be reinfected, the rate of change of the recovered population is

$$\dot{R}_t = \alpha I_t - \mu R_t, \tag{4}$$

the inflow of newly recovered minus the outflow from the stock of recovered individuals who die of natural causes.

[6] Given the Poisson structure of the model, and hence the stationarity of consumers' life cycles, assuming vaccines are adminstered only to newborns is without loss of generality; we could equivalently have assumed that the vaccine is administered to any subset of susceptible consumers who have not yet been vaccinated.

Table 1
Estimates of $\mathscr{R}_0$ from the epidemiology literature.

Disease	Time period	Location	$\mathscr{R}_0$	Source
SARS	2002	Singapore	1.1	Chowell et al. (2003)
Covid-19 (ancestral)	2020	Global average	2.8	Liu et al. (2020)
HIV	1981–85	United Kingdom	2–5	Anderson and May (1991)
Covid-19 (Delta)	2021	Global average	5.1	Liu and Rocklöv (2021)
Rubella	1960–70	United Kingdom	6–7	Anderson and May (1991)
Chicken pox	1944–68	United Kingdom	10–12	Anderson and May (1991)
Mumps	1960–80	United Kingdom	11–14	Anderson and May (1991)
Measles	1950–68	United Kingdom	16–18	Anderson and May (1991)
Pertussis	1944–78	United Kingdom	16–18	Anderson and May (1991)

Notes: All Anderson and May (1991) estimates are from their Table 4.1, reporting estimates from studies of England and Wales within the United Kingdom. The HIV estimate is for Type 1 in the subpopulation of men who have sex with men (MSM).

Newborns enter the population as susceptibles. The birth rate equals the mortality rate, μ, maintaining a stationary population size. A fraction θQ of newborns are immunized at birth. Mass $\beta I_t S_t$ of susceptibles become infected each instant, and a further μS_t die. Thus, the rate of change of the susceptible population is

$$\dot{S}_t = \mu - \theta Q - \beta I_t S_t - \mu S_t. \tag{5}$$

In lieu of the transmission parameter β, epidemiologists often work with a related parameter called the basic reproductive number $\mathscr{R}_0$, equal to the expected number of secondary cases an infectious individual would transmit in a fully susceptible population.[7] One can see that the disease eventually dies out in an unvaccinated population if $\mathscr{R}_0 < 1$ and remains endemic if $\mathscr{R}_0 > 1$. In our model,

$$\mathscr{R}_0 = \frac{\beta}{\alpha + \mu}. \tag{6}$$

To provide intuition for Eq. (6), each instant an individual remains infected, he or she infects a number of others equal to β times the size of the susceptible population—approximately 1 over the period since the infected individual is introduced into a fully susceptible population. Owing to the Poisson structure of the model, the individual remains infected for an expected duration of $1/(\alpha + \mu)$.[8]

The subsequent analysis takes $\mathscr{R}_0$ as the key exogenous parameter, capturing the disease's infectiousness. Table 1 presents estimates of $\mathscr{R}_0$ from the epidemiology literature, varying from 1.1 for SARS at the low end to 16–18 for measles and pertussis at the high end. Since estimates of $\mathscr{R}_0$ can vary across time and region if his or her marginal private bene as well as across disease, the table lists the relevant time period and region from which each estimate comes.

2.2. Steady state

Denote steady-state values of compartment sizes by dropping their time subscripts and appending Q as an argument to emphasize their dependence on that key variable to be endogenized later. We can compute these steady-state values by solving the system of equations formed by setting $\dot{S}_t = \dot{I}_t = \dot{R}_t = \dot{Z}_t = 0$ in Eqs. (2)–(5).

Focusing on the infected compartment, which is the most important in the subsequent analysis, the unique stable solution can be written

$$I(Q) = \begin{cases} 0 & Q \geq \mathscr{Q}_0 \\ \dfrac{\mu}{\alpha + \mu}\left(1 - \dfrac{1}{\mathscr{R}_0} - \dfrac{\theta}{\mu}Q\right) & Q < \mathscr{Q}_0, \end{cases} \tag{7}$$

where

$$\mathscr{Q}_0 = \max\left[0, \frac{\mu}{\theta}\left(1 - \frac{1}{\mathscr{R}_0}\right)\right] \tag{8}$$

is the threshold vaccine quantity above which the disease dies out and below which the disease remains endemic.[9]

Gans (2020) argues that a useful shorthand for capturing rational consumer distancing in an epidemiological model is to constrain the effective equilibrium number to equal 1. We can show that this is an inherent property of a nontrivial steady

[7] The modern definition of $\mathscr{R}_0$ due to Diekmann et al. (1990) is the dominant eigenvalue of the next-generation operator in the epidemiological system. Martcheva (2015, p. 51) shows that Eq. (6) provides the value of $\mathscr{R}_0$ implied by this definition.

[8] To see this, note that there are two competing risks for exiting the infected state: recovery, which has hazard $\lambda_R(t) = \alpha$, and mortality, which has hazard $\lambda_M(t) = \mu$. The combined hazard of exiting the infected state is $\lambda_{EI}(t) = \lambda_R(t) + \lambda_M(t)$. As is well known for Poisson duration models, the duration of a spell equals the reciprocal of the hazard, here $1/\lambda_{EI}(t) = 1/(\alpha + \mu)$.

[9] The trivial solution $I(Q) = 0$ always exists, but it is unstable when $\mathscr{R}_0 > \mu/(\mu - \theta Q)$.

state of our model using the concept of the effective reproductive number, $\mathscr{R}_t$. The effective reproductive number, given by the formula $\mathscr{R}_t = \mathscr{R}_0 S_t$, equals the number of secondary cases an infectious individual transmits in the observed population. Denote the steady-state value of the effective reproductive number by $\mathscr{R}(Q) = \mathscr{R}_0 S(Q)$. Suppose the steady state involves a nontrivial infection rate, $I(Q) > 0$. In that case, for Eq. (3) to equal 0—as required in a steady state—$\beta S(Q) = \alpha + \mu$, implying $\mathscr{R}_0 S(Q) = 1$ after substituting from (6), in turn implying $\mathscr{R}(Q) = 1$, as claimed.

2.3. Consumer demand

Having specified the epidemiological features of the model, we turn to economic features of the vaccine market, starting with demand. Assume consumers are risk neutral. For now, assume consumers are homogeneous; Appendix C extends the results to heterogeneous consumers. To avoid computing transition paths, focus on the steady-state in the limiting case without discounting.

A consumer demands the vaccine if his or her marginal private benefit $\text{MPB}(Q)$ from the vaccine exceeds its price P. For a risk-neutral consumer, $\text{MPB}(Q) = \theta H \Phi(Q)$, where efficacy parameter θ is the probability that the vaccine allows the consumer to avoid harm, H denotes the expected harm over the infection's spell conditional on being infected, and $\Phi(Q)$ denotes the probability that a susceptible person becomes infected.

To compute H, harm measured as a stock over the infection spell, start by defining h to be flow harm to a consumer each instant he or she is infected. The consumer experiences flow harm over the infection spell, which we showed has expected duration $1/(\alpha + \mu)$ (see footnote 8). Hence,

$$H = \frac{h}{\alpha + \mu}. \tag{9}$$

More involved calculations can be used to show the probability that a susceptible person becomes infected equals[10]

$$\Phi(Q) = \begin{cases} 0 & Q \geq \mathscr{Q}_0 \\ 1 - \dfrac{1}{(1 - \theta Q/\mu)\mathscr{R}_0} & Q < \mathscr{Q}_0. \end{cases} \tag{10}$$

The marginal private benefit inherits this branched structure from $\Phi(Q)$:

$$\text{MPB}(Q) = \begin{cases} 0 & Q \geq \mathscr{Q}_0 \\ \theta H\left[1 - \dfrac{1}{(1 - \theta Q/\mu)\mathscr{R}_0}\right] & Q < \mathscr{Q}_0. \end{cases} \tag{11}$$

Proceeding to derive the demand curve, all μ newborns who are eligible to purchase the vaccine do so if $P < \text{MPB}(\mu)$, and none purchase if $P > \text{MPB}(0)$. For P strictly between $\text{MPB}(\mu)$ and $\text{MPB}(0)$, some but not all consumers purchase. Given they are homogeneous, consumers must be indifferent between purchasing and not, implying $P = \text{MPB}(Q)$. Given they are indifferent, any fraction of them are willing to purchase in equilibrium. Demand is pinned down by the value of Q satisfying (11) when the right-hand side is set equal to P. Demand in that case is given by

$$d(P) = \frac{\mu}{\theta}\left[1 - \frac{1}{(1 - P/\theta H)\mathscr{R}_0}\right]. \tag{12}$$

Combining these facts yields the demand curve

$$D(P) = \begin{cases} 0 & P > \text{MPB}(0) \\ d(P) & P \in [\text{MPB}(\mu), \text{MPB}(0)] \\ \mu & P < \text{MPB}(\mu). \end{cases} \tag{13}$$

Equivalently, the demand curve is given by $d(P)$ unless this violates a boundary condition $d(P) \in [0, \mu]$, in which case demand is given by the violated boundary.

Demand considerations alone allow us to establish that Geoffard and Philipson's (1997) important result—nontrivial equilibria cannot entail the disease's eradication in the steady state—holds quite generally in our model as well. The result holds independent of market structure and holds whether or not the government subsidizes vaccines.

[10] An individual has two competing risks for exiting the susceptible state: infection, which has hazard $\lambda_I(t, Q) = \beta I(Q) = (\alpha + \mu)\mathscr{R}_0 I(Q)$ each instant, and mortality from natural causes, which has hazard $\lambda_M(t) = \mu$. The combined hazard of exiting the susceptible state is $\lambda_C(t, Q) = \lambda_I(t, Q) + \lambda_M(t)$. By standard results for competing Poisson risks, the consumer's cumulative risk of exiting the susceptible state by age t is $\Lambda_C(t, Q) = \int_0^t \lambda_C(\tau, Q)d\tau$, probability of surviving as susceptible to age t is $e^{-\Lambda_C(t,Q)}$, and likelihood of exit due to infection is $\phi(t, Q) = \lambda_I(t, Q)e^{-\Lambda_C(t,Q)}$. To compute the probability that the consumer experiences an infection at some point over his or her lifetime, we integrate this cause-specific likelihood:

$$\Phi(Q) = \int_0^\infty \phi(t, Q)dt = \frac{(\alpha + \mu)\mathscr{R}_0 I(Q)}{(\alpha + \mu)\mathscr{R}_0 I(Q) + \mu}.$$

Substituting for $I(Q)$ from (7) yields Eq. (10).

Proposition 1. *Consider an arbitrary vaccine market structure with or without government subsidies. If the disease is eradicated in steady-state equilibrium in this market ($I^* = 0$), then either the vaccine is free to consumers ($P^* = 0$) or no vaccine is purchased ($Q^* = 0$).*

The proof is provided in Appendix A. Intuitively, if the disease is eradicated in steady-state equilibrium, then $\Phi(Q^*) = 0$ by (10), implying $\text{MPB}(Q^*) = \theta H \Phi(Q^*) = 0$. But consumers with no marginal private benefit will not purchase the vaccine at a positive price.

2.4. Firm supply

The main text of this paper focuses on two polar opposites for market structure on the supply side: perfect competition and monopoly. We also analyze a model of Cournot competition among n firms. We relegate the relatively complex details involved in the Cournot analysis to Appendix B, briefly summarizing the results in the text.

Firms are assumed to produce at constant marginal and average cost $c > 0$ per vaccine course (where a course is the approved regimen, possibly involving multiple doses). Under perfect competition, vaccine supply is perfectly elastic at price c. Under monopoly, the firm chooses a price maximizing the flow of industry profit Π from steady-state sales.[11]

2.5. Normative measures

Define aggregate health in the steady state as

$$\text{AH}(Q) = h[1 - I(Q)], \tag{14}$$

the product of an individual's avoided harm h and the number of healthy individuals $1 - I(Q)$ at any instant in steady-state equilibrium. The social benefit of the vaccine is the difference between aggregate health with and without the vaccine: $\text{SB}(Q) = \text{AH}(Q) - \text{AH}(0) = h[I(0) - I(Q)]$. Substituting from (7) and (9) and condensing cases using (8) yields

$$\text{SB}(Q) = \theta H \min(Q, \mathscr{Q}_0). \tag{15}$$

Let $\text{MSB}(Q) = \partial \text{SB}(Q)/\partial Q$ denote marginal social benefit from a vaccine course. Differentiating (15),[12]

$$\text{MSB}(Q) = \begin{cases} 0 & Q \geq \mathscr{Q}_0 \\ \theta H & Q < \mathscr{Q}_0. \end{cases} \tag{16}$$

Let $\text{MEX}(Q) = \text{MSB}(Q) - \text{MPB}(Q)$ denote the marginal externality from a vaccine course. Subtracting the formulas provided yields

$$\text{MEX}(Q) = \begin{cases} 0 & Q \geq \mathscr{Q}_0 \\ \dfrac{\theta H}{(1 - \theta Q/\mu)\mathscr{R}_0} & Q < \mathscr{Q}_0. \end{cases} \tag{17}$$

Let $W(Q)$ denote social welfare in the market, aggregate health minus total vaccine production costs:

$$W(Q) = \text{AH}(Q) - cQ. \tag{18}$$

A comprehensive analysis of these normative measures is postponed until after we have solved for the equilibrium value of Q, on which the measures depend. Some observations can be made even at this preliminary stage. For $Q \geq \mathscr{Q}_0$, enough vaccine is available to eradicate the disease in the steady state, eliminating the marginal benefit of vaccine. Hence, $\text{MPB}(Q) = \text{MSB}(Q) = \text{MEX}(Q) = 0$ for $Q \geq \mathscr{Q}_0$.

For $Q < \mathscr{Q}_0$, $\text{MSB}(Q) = \theta H$, which is the expected benefit of treating an infected person with a drug that does not prevent disease transmission but has efficacy θ in preventing harmful symptoms. It may be at first puzzling that vaccinating an individual not certain to contract the disease amounts to the same social benefit as treating a certainly infected individual. The puzzle is resolved by recognizing the vaccine's positive externality. Conditional on the vaccine being effective for the individual, an event with probability θ, his or her private benefit plus the chain of expected external benefits sum exactly to H in the model.[13]

Also remarkable is the fact that, for $Q < \mathscr{Q}_0$, $\text{MSB}(Q)$ is a constant that does not vary with the infectiousness of the disease as measured by $\mathscr{R}_0$. When $\mathscr{R}_0$ increases, the vaccinated individual's direct benefit $\text{MPB}(Q)$ increases; but, perhaps

[11] A technical issue arises with the definition of price in equilibria in which output is zero. These equilibria arise under all market structures for $\mathscr{R}_0$ sufficiently low that consumers' willingness to pay is below c. Equilibrium price is not unique then: an interval of prices from below c to infinity are consistent with zero output. Under perfect competition, the obvious convention for equilibrium price is simply to define it to be c whether or not output is zero. Less obvious is the appropriate convention when firms have market power; for concreteness, we will define equilibrium price to be c when output is zero regardless of market structure.

[12] The derivative $\partial \text{SB}(\mathscr{Q}_0)\partial Q$ does not exist, but the left and right derivatives, respectively $\partial \text{SB}(\mathscr{Q}_0)/\partial Q^-$ and $\partial \text{SB}(\mathscr{Q}_0)/\partial Q^+$, do. We set $\text{MSB}(\mathscr{Q}_0) = \partial \text{SB}(\mathscr{Q}_0)/\partial Q^+$.

[13] In more complicated epidemiological models, $\text{MSB}(Q)$ need not simply be a constant. That it is a constant in our version of the SIR model hinges on the simple specification of new infections, $\beta I_t S_t$, in Eq. (3).

Table 2
Steady-state equilibrium variables under perfect competition as functions of $\mathscr{R}_0$.

Variable	Case (a) $\mathscr{R}_0 \in [0, 1]$	Case (b) $\mathscr{R}_0 \in \left(1, \frac{1}{1-\tilde{c}}\right]$	Case (c) $\mathscr{R}_0 \in \left(\frac{1}{1-\tilde{c}}, \frac{1}{(1-\theta)(1-\tilde{c})}\right]$	Cases (d) and (e) $\mathscr{R}_0 \in \left(\frac{1}{(1-\theta)(1-\tilde{c})}, \infty\right)$
P_c^*	c	c	c	c
Q_c^*	0	0	$\frac{\mu}{\theta}\left[1 - \frac{1}{(1-\tilde{c})\mathscr{R}_0}\right]$	μ
Π_c^*	0	0	0	0
I_c^*	0	$\frac{\mu}{\alpha+\mu}\left(1 - \frac{1}{\mathscr{R}_0}\right)$	$\frac{\mu}{\alpha+\mu}\left[\frac{\tilde{c}}{(1-\tilde{c})\mathscr{R}_0}\right]$	$\frac{\mu}{\alpha+\mu}\left(1 - \theta - \frac{1}{\mathscr{R}_0}\right)$
MPB_c^*	0	$\theta H\left(1 - \frac{1}{\mathscr{R}_0}\right)$	c	$\theta H\left[1 - \frac{1}{(1-\theta)\mathscr{R}_0}\right]$
MSB_c^*	0	θH	θH	θH
MEX_c^*	0	$\frac{\theta H}{\mathscr{R}_0}$	$\theta H(1-\tilde{c})$	$\frac{\theta H}{(1-\theta)\mathscr{R}_0}$
W_c^*	$H(\alpha+\mu)$	$H\left(\alpha + \frac{\mu}{\mathscr{R}_0}\right)$	$H[\alpha + \mu(1-\tilde{c})]$	$H\left[\alpha + \theta\mu(1-\tilde{c}) + \frac{\mu}{\mathscr{R}_0}\right]$

Notes: The distinction between cases (d) and (e) in the last column, relevant for monopoly in the next table, is irrelevant for perfect competition here.

counterintuitively, the marginal externality $\text{MEX}(Q)$ decreases. An individual's vaccination provides less external benefit to contacts because they are increasingly likely to pick up the disease from someone else. The increase in $\text{MPB}(Q)$ and decrease in $\text{MEX}(Q)$ exactly offset so that their sum $\text{MSB}(Q)$ is independent of $\mathscr{R}_0$ in the model.[14]

As a benchmark, we will characterize the first-best quantity Q^{**}, which maximizes social welfare. To accommodate cases in which a possibly open set of quantities maximizes social welfare, we take Q^{**} to be the infimum of the set. Comparing the marginal social benefit of a vaccine in (16) (which we see is either 0 or a positive constant) with its marginal social cost (given simply by c), it is immediate that $Q^{**} = 0$ if $c \geq \theta H$. To rule out this trivial case, we assume throughout the remainder of the paper that the disease is harmful enough that the vaccine's social benefit exceeds its cost: $\theta H > c$. Rescaling cost as $\tilde{c} = c/\theta H$, the assumption can be equivalently written

$$\tilde{c} < 1. \tag{19}$$

When (19) holds, we can see from (16) that $Q^{**} = \mathscr{Q}_0$ unless this quantity exceeds the flow of newborns, indicating that the first best is a corner solution in which all newborns are vaccinated: $Q^{**} = \mu$. Thus,

$$Q^{**} = \min(\mathscr{Q}_0, \mu). \tag{20}$$

3. Equilibrium

3.1. Perfect competition

Equilibrium values of steady-state variables will be distinguished with asterisks, with an added subscript indicating the relevant market structure. Thus, for example, P_c^* denotes the perfectly competitive equilibrium price.

Under perfect competition, $P_c^* = c$. The remaining equilibrium variables can be computed using straightforward algebra applied to the supplied formulas taking due care to ensure the appropriate branch of each formula is used in the relevant case. The relevant cases turn out to depend on the interval in which $\mathscr{R}_0$ falls.

First, consider $\mathscr{R}_0 \leq 1$. Infectiousness is so low in this case—labeled (a)—that the disease disappears in the steady state even without a vaccine. The vaccine has no marginal social or private value, and no vaccine is sold.

For values of $\mathscr{R}_0$ just above 1, the disease is infectious enough not to disappear in the steady state but not infectious enough to justify its purchase at the competitive price. Although the vaccine has a positive marginal social and private benefit, the marginal private benefit is below c even when no other consumer purchases. The upper bound on $\mathscr{R}_0$ for this case—labeled (b)—is determined by the condition $\text{MPB}(0) \leq c$, which using (11) yields $\mathscr{R}_0 \leq 1/(1-\tilde{c})$.

For $\mathscr{R}_0$ in an intermediate range—labeled case (c)—some but not all newborns purchase. The interior solution for equilibrium quantity Q_c^* can be found by substituting $P_c^* = c$ into demand expression (12). The values of other equilibrium variables can be found by substituting Q_c^* into the relevant equations for them.

[14] That an increase in $\mathscr{R}_0$ causes opposing changes in $\text{MPB}(Q)$ and $\text{MEX}(Q)$ is robust across epidemiological models. That they exactly offset hinges on the simple structure of our version of the SIR model as discussed in the previous footnote.

Table 3
Steady-state equilibrium variables under monopoly as functions of $\mathscr{R}_0$.

Variable	Case (a) $\mathscr{R}_0 \in [0, 1]$	Case (b) $\mathscr{R}_0 \in \left(1, \frac{1}{1-\tilde{c}}\right]$	Cases (c) and (d) $\mathscr{R}_0 \in \left(\frac{1}{1-\tilde{c}}, \frac{1}{(1-\theta)^2(1-\tilde{c})}\right]$	Case (e) $\mathscr{R}_0 \in \left(\frac{1}{(1-\theta)^2(1-\tilde{c})}, \infty\right)$
P_m^*	c	c	$\theta H\left(1-\sqrt{\frac{1-\tilde{c}}{\mathscr{R}_0}}\right)$	$\theta H\left[1-\frac{1}{(1-\theta)\mathscr{R}_0}\right]$
Q_m^*	0	0	$\frac{\mu}{\theta}\left[1-\frac{1}{\sqrt{(1-\tilde{c})\mathscr{R}_0}}\right]$	μ
Π_m^*	0	0	$\frac{\mu H(1-\tilde{c})}{\theta}\left[1-\frac{1}{\sqrt{(1-\tilde{c})\mathscr{R}_0}}\right]^2$	$\mu\theta H\left[1-\tilde{c}-\frac{1}{(1-\theta)\mathscr{R}_0}\right]$
I_m^*	0	$\frac{\mu}{\alpha+\mu}\left(1-\frac{1}{\mathscr{R}_0}\right)$	$\frac{\mu}{\alpha+\mu}\left[\frac{1}{\sqrt{(1-\tilde{c})\mathscr{R}_0}}-\frac{1}{\mathscr{R}_0}\right]$	$\frac{\mu}{\alpha+\mu}\left(1-\theta-\frac{1}{\mathscr{R}_0}\right)$
MPB_m^*	0	$\theta H\left(1-\frac{1}{\mathscr{R}_0}\right)$	$\theta H\left(1-\sqrt{\frac{1-\tilde{c}}{\mathscr{R}_0}}\right)$	$\theta H\left[1-\frac{1}{(1-\theta)\mathscr{R}_0}\right]$
MSB_m^*	0	θH	θH	θH
MEX_m^*	0	$\frac{\theta H}{\mathscr{R}_0}$	$\theta H\sqrt{\frac{1-\tilde{c}}{\mathscr{R}_0}}$	$\frac{\theta H}{(1-\theta)\mathscr{R}_0}$
W_m^*	$H(\alpha+\mu)$	$H\left(\alpha+\frac{\mu}{\mathscr{R}_0}\right)$	$H\left[\alpha+\mu\left(1-\tilde{c}+\frac{1}{\mathscr{R}_0}-\sqrt{\frac{1-\tilde{c}}{\mathscr{R}_0}}\right)\right]$	$H\left[\alpha+\mu\theta(1-\tilde{c})+\frac{\mu}{\mathscr{R}_0}\right]$

Notes: The distinction between cases (c) and (d), relevant for perfect competition in the previous table, is irrelevant for monopoly here. In cases (a) and (b), in which equilibrium output is zero, any price in an interval from below c to infinity is consistent with zero output. As discussed in footnote 11, we adopt the convention that $P_m^* = c$ in those cases.

For extreme values of $\mathscr{R}_0$, the disease can conceivably be so infectious that all μ newborns purchase at the competitive price. The existence of this case requires an imperfectly effective vaccine, $\theta < 1$. Vaccinating all consumers with a perfectly effective vaccine would eliminate the infection and demand, which would be inconsistent with all newborns purchasing. Enough unsuccessfully vaccinated consumers must remain to generate an infection rate that justifies purchase at the competitive price. The relevant condition determining this case—labeled (d)—is $\text{MPB}(\mu) \geq c$, implying $\mathscr{R}_0 \geq 1/(1-\theta)(1-\tilde{c})$. Since all consumers purchase in this case, there is no underconsumption distortion; the first best is obtained.

Table 2 reports the steady-state equilibrium values of selected variables under perfect competition as a function of $\mathscr{R}_0$, and a selection of these are graphed in Fig. 1. Vaccine quantity Q_c^*, the dashed curve in the first panel, equals 0 in case (a) because there is no infection in equilibrium. It continues to be 0 in case (b) despite positive disease prevalence because the prevalence is too low for any consumer to justify spending c on it. In case (c), we start to see a positive vaccine quantity, which increases in $\mathscr{R}_0$ until case (d) is reached and all μ purchase.

As might be expected, Q_c^*, MPB_c^*, and W_c^* are monotonic in the infectiousness of the disease measured by $\mathscr{R}_0$, with Q_c^* weakly increasing, MPB_c^* weakly increasing, and W_c^* weakly decreasing. Other variables are nonmonotonic. Disease prevalence I_c^* is increasing in case (b) due to the direct epidemiological effect of the higher $\mathscr{R}_0$ with no vaccine purchases in that case to offset it. In case (c), I_c^* reverses course and begins to decrease in $\mathscr{R}_0$ in an epidemiological version of the Peltzman (1975) effect, whereby consumers' risk-compensation activity, in this case increased vaccination, offsets the environmental increase in risk. In cases (d) and (e), I_c^* again rises with $\mathscr{R}_0$. The direct effect of an increase in infectiousness cannot be offset by an increase in Q_c^* given that all newborns are already purchasing in these cases.

The marginal externality MEX_c^* also exhibits an interesting nonmonotonicity. Although MEX_c^* is monotonic over the subinterval ($\mathscr{R}_0 > 1$) for which it is positive, it is nonmonotonic over the entire range of $\mathscr{R}_0$: MEX_c^* starts out at 0 for $\mathscr{R}_0 < 1$, exhibits a discontinuous jump up to its global peak for $\mathscr{R}_0$ just above 1, and declines for yet higher levels of $\mathscr{R}_0$ than 1. For $\mathscr{R}_0 > 1$, the greater infectiousness of the disease leads consumers to internalize an increasing share of the social benefit of their own vaccination.

The next proposition, proved in Appendix A, provides a general summary of the comparative-static effects.

Proposition 2. *Consider the comparative-static effect of $\mathscr{R}_0$ on steady-state equilibrium under perfect competition.*

- *Price and industry profit are constant for $\mathscr{R}_0 > 0$, with $P_c^* = c$ and $\Pi_c^* = 0$.*
- *Q_c^* and MPB_c^* are weakly increasing in $\mathscr{R}_0$ for all $\mathscr{R}_0 > 0$.*
- *I_c^* has a local interior peak $I_c^* = \mu\tilde{c}/(\alpha+\mu)$ at $\mathscr{R}_0 = 1/(1-\tilde{c})$. This local maximum is a global maximum if and only if $\theta \geq 1-\tilde{c}$.*
- *Marginal social benefit is constant for all $\mathscr{R}_0 > 1$, with $\text{MSB}_c^* = \theta H$.*
- *The marginal externality is nonmonotonic over $\mathscr{R}_0 > 0$, approaching an interior supremum $\text{MEX}_c^* = \theta H$ as $\mathscr{R}_0 \downarrow 1$.*

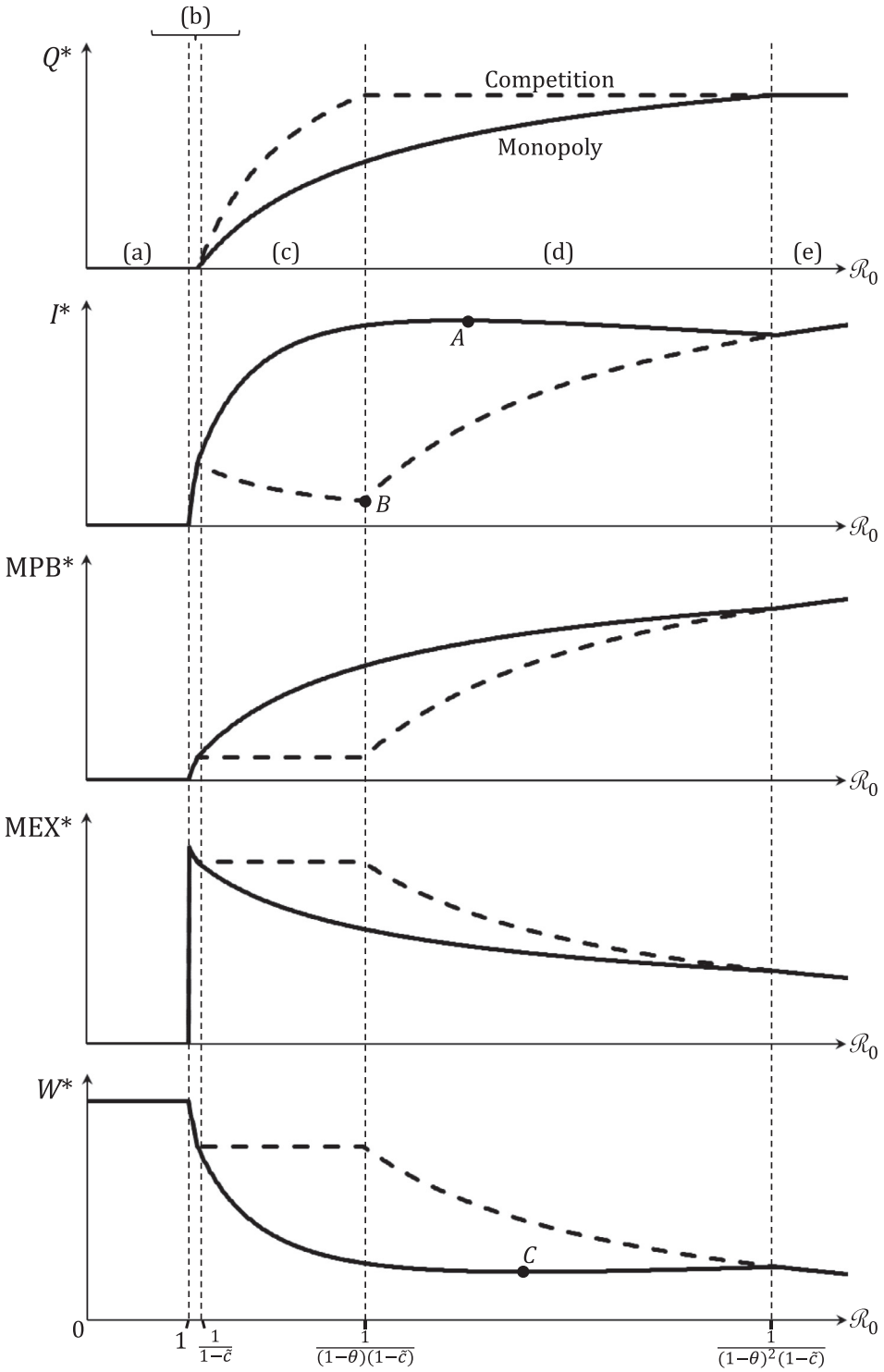

Fig. 1. Graphs of Steady-State Equilibrium Variables as Functions of $\mathscr{R}_0$. *Notes:* Graph of formulas provided in Tables 2 and 3, illustrated for specific parameter values ($\theta = 0.6$, $\alpha = 0$, $\mu = 1$,$c = 0.1$, $H = 1$). Dashed curves represent equilibrium under perfect competition and solid curves under monopoly. Where dashed and solid curves overlap, the solid curve represents both industry structures.

- W_c^* *is weakly decreasing in* $\mathscr{R}_0$ *for all* $\mathscr{R}_0 > 0$.

The weak changes in Q_c^*, MPB_c^*, *and* W_c^* *are strict for a nonempty interval of* $\mathscr{R}_0$ *for each variable.*

3.2. Monopoly

Output is never higher under monopoly than perfect competition, implying $Q_m^* = 0$ in cases (a) and (b) in which $Q_c^* = 0$. In the remaining cases, competitive firms sell a positive amount at price c. By continuity, the monopoly can sell a positive amount at a small markup above c, earning a positive profit, implying $Q_m^* > 0$ for $\mathscr{R}_0$ in case (c) and above.

To solve for Q_m^* in these other cases, it is convenient to express the monopoly's maximization problem so that the choice variable is quantity rather than price. Letting $P(Q)$ denote inverse demand, monopoly profit can be written $\Pi(Q) = [P(Q) - c]Q = [\text{MPB}(Q) - c]Q$, where the second equality follows since the monopoly optimally charges a price extracting the entire private benefit of the marginal consumer. The monopoly maximizes profit subject to the constraint that no more than the population of newborn consumers can be served: $Q \leq \mu$. Applying the Kuhn-Tucker method yields the following solution. If $\mathscr{R}_0 < 1/(1-\theta)^2(1-\tilde{c})$, then the constraint $Q \leq \mu$ does not bind, yielding solution

$$Q_m^* = \frac{\mu}{\theta}\left[1 - \sqrt{\frac{1}{(1-\tilde{c})\mathscr{R}_0}}\right]. \tag{21}$$

Otherwise, the constraint holds with equality, implying $Q_m^* = \mu$. A necessary condition for the constraint to bind is that the vaccine is imperfectly effective ($\theta < 1$). The monopoly would never sell a perfectly effective vaccine to all consumers because this would eradicate the disease, leaving the monopoly with no steady-state profit according to Proposition 1.

Substituting Q_m^* into the formulas supplied for the other variables yields the entries in Table 3. The solid curves in Fig. 1 help visualize how the equilibrium monopoly variables in Table 3 vary with $\mathscr{R}_0$ and how the monopoly compares with perfect competition shown as the dashed curves. The two market structures overlap in cases (a) and (b), neither generating any vaccine output. The two market structures overlap again in (e), both generating the first-best quantity $Q^{**} = \mu$. In between—in (c) and (d)—the two market structures diverge, with monopoly generating strictly lower output. The strictly lower output under monopoly in these cases leads to a weakly higher infection rate and marginal private benefit and weakly lower welfare under monopoly. Perhaps counterintuitively, the marginal externality is also lower under monopoly. This follows from the fact that $\text{MPB}_m^* \geq \text{MPB}_c^*$, implying $\text{MEX}_m^* = \text{MSB}_m^* - \text{MPB}_m^* \leq \text{MSB}_c^* - \text{MPB}_c^* = \text{MEX}_c^*$ since $\text{MSB}_m^* = \text{MSB}_c^* = \theta H$.

While it is not surprising that welfare is lower under monopoly than competition, as this is true in typical markets, our model shuts down the typical channel for monopoly deadweight loss by taking consumers to be homogeneous. In our model, the epidemiological externality confers market power: starting from a price that extracts purchasers' entire marginal private benefit, a positive albeit reduced fraction of consumers will continue to purchase at a higher price since the reduction in vaccine quantity increases their marginal private benefit through an increase in disease prevalance. The monopoly's exercise of this market power generates deadweight loss. The large gap between W_c^* and W_m^* for an intermediate range of $\mathscr{R}_0$ in the bottom panel of Fig. 1 suggests that monopoly distortions may be worst for moderate levels of infectiousness. Market power generates little welfare loss for the lowest values of $\mathscr{R}_0$ and—with an imperfect vaccine—for the highest values of $\mathscr{R}_0$.

The comparative statics for Q^*, MPB^*, and MEX^* are similar across the two market structures; but the comparative statics for I^* are more complex for monopoly than perfect competition. The direct, epidemiological effect leading an increase in infectiousness $\mathscr{R}_0$ to increase prevalence I_m^* is still present. So is the opposing risk-compensation effect, leading consumers to increase vaccine purchases. However, the monopoly can leverage its market power to absorb some of consumers' increased willingness to pay to avoid risk by increasing price, dampening the risk-compensation effect. Depending on the parameters, the balance between forces need not tip in the same direction as under perfect competition.

The graph of W^* under monopoly illustrates the remarkable possibility that increasing $\mathscr{R}_0$ can increase welfare. One would think that society would always be harmed by an increase in infectiousness. While the direct, epidemiological effect of an increase in $\mathscr{R}_0$ harms society, the indirect effect of overcoming the twin distortions of consumer free riding and monopoly market power can more than offset the direct effect, increasing welfare over some parameter ranges. In the bottom panel of Fig. 1, we see this possibility emerging for the interval of $\mathscr{R}_0$ between point C and the boundary of (d). Under monopoly, not only do consumers fail to consider the external benefit their vaccination provides other consumers, but the monopoly compounds this by placing negative value on consumption to the extent it reduces others' willingness to pay for a vaccine. Mitigating this compounded underconsumption problem via an increase in $\mathscr{R}_0$ can provide such a large indirect benefit that it swamps the direct harm from an increase in $\mathscr{R}_0$, leading to an increase in social welfare. The possibility that welfare increases with $\mathscr{R}_0$ never arises under perfect competition because it requires the distortion due to supplier market power to compound the free-rider problem.

The next proposition, proved in Appendix A, summarizes the comparative-static effects of an increase in $\mathscr{R}_0$ on the steady-state equilibrium under monopoly observed from Fig. 1.

Proposition 3. *Consider the comparative-static effect of* $\mathscr{R}_0$ *on steady-state equilibrium under monopoly.*

- Q_m^*, P_m^*, MPB_m^*, *and* Π_m^* *are weakly increasing in* $\mathscr{R}_0$ *for all* $\mathscr{R}_0 > 0$.
- *If* θ *and* $\tilde{c}$ *are sufficiently low, then* I_m^* *is weakly increasing in* $\mathscr{R}_0$ *for all* $\mathscr{R}_0 > 0$. *Otherwise,* I_m^* *has an interior peak, which is a global maximum for sufficiently high* θ.

- *Marginal social benefit is constant for all* $\mathscr{R}_0 > 1$, *with* $\text{MSB}_m^* = \theta H$.
- *The marginal externality has a global peak for an interior value of* $\mathscr{R}_0$. *In particular, the supremum* $\text{MEX}_m^* = \theta H$ *is approached in the limit* $\mathscr{R}_0 \downarrow 1$.
- *If* θ *is sufficiently low, then* W_m^* *is weakly decreasing in* $\mathscr{R}_0$ *for all* $\mathscr{R}_0 > 0$. *Otherwise,* W_m^* *is nonmonotonic in* $\mathscr{R}_0$, *reaching a local interior minimum, which is a global minimum for sufficiently high* θ.

The weak changes in Q_c^*, MPB_c^*, Π_m^*, *and* P_m^* *are strict for a nonempty interval of* $\mathscr{R}_0$ *for each variable.*

Proposition 3 provides qualitative conditions under which I_m^* and W_m^* are nonmonotonic and achieve global interior optima. The proof of Proposition 3 provides explicit conditions for these outcomes as well as closed-form expressions for the location and value of the local optima. Using those explicit conditions, we can derive simple sufficient conditions for I_m^* and W_m^* to achieve local and global interior optima. Those sufficient conditions are stated in the next proposition, proved in Appendix A.

Proposition 4. *If* $\theta > 1/2$, I_m^* *reaches a local maximum and* W_m^* *a local minimum over* $\mathscr{R}_0 > 0$. *These local optima are global optima if* $\theta > 3/4$.

According to the proposition, if the vaccine is more than 75% effective, then I_m^* and W_m^* attain global optima for interior values of $\mathscr{R}_0$. In practice, agencies such as the U.S. Food and Drug Administration (FDA) do not typically license vaccines with efficacy below 80% (Brennan, 2009). For realistic parameter values we therefore have the surprising result that if a vaccine is sold by a monopoly directly to consumers, disease prevalence will be greatest and social welfare lowest not for diseases with the most extreme infectiousness as indexed by $\mathscr{R}_0$ but for moderate infectiousness.[15]

3.3. Cournot competition

Appendix B provides an analysis of Cournot competition among n firms, complete with a table of equilibrium variables for various values of $\mathscr{R}_0$ analogous to Table 3. The appendix shows that the comparative-statics results under monopoly in Proposition 3 hold verbatim under Cournot competition for all $n \geq 1$. The appendix also shows that the effects of increasing n that one would expect to hold in a Cournot model do hold: for any $\mathscr{R}_0$ for which variables vary with n, increasing n reduces equilibrium price, profit, and infections and increases quantity and welfare.

4. Government subsidies

We have seen that positive externalities can lead to inefficiently low vaccine consumption relative to the first best. This section characterizes the optimal government subsidy to correct this distortion and determines its comparative-static properties.

Assume that the government has lexicographic preferences, maximizing social welfare as its primary goal, breaking ties with the secondary goal of minimizing expenditures. The government commits to a per-course subsidy $G \geq 0$ paid to firms, equivalent to a reduction in their marginal cost from c to $c - G$. Since the first-best quantity Q^{**} maximizes social welfare, the first-best subsidy G^{**} is that implementing Q^{**}. If a set of subsidies implements Q^{**}, the government's secondary objective is fulfilled by taking G^{**} to be the infimum of the set.

Before analyzing particular market structures, we provide some general principles behind the optimal subsidy that apply to any market structure.

Proposition 5. *Consider any market structure in which an increase in a subsidy weakly reduces the equilibrium price. If* $\mathscr{R}_0 \leq 1$, *then* $G^{**} = 0$; *the disease is eradicated in the steady state without a vaccine or subsidy. If* $\mathscr{R}_0 \in (1, 1/(1-\theta))$, *then* G^{**} *induces equilibrium quantity* $Q^{**} = \mathscr{Q}_0$ *and price* $P^{**} = 0$, *resulting in eradication of the disease in the steady state. If* $\mathscr{R}_0 > 1/(1-\theta)$, *then* G^{**} *induces equilibrium quantity* $Q^{**} = \mu$ *and price*

$$P^{**} = \theta H\left[1 - \frac{1}{(1-\theta)\mathscr{R}_0}\right], \tag{22}$$

resulting in universal vaccination but not eradication.

Appendix A provides a proof. Intuitively, the government would like to eradicate the disease in all circumstances if this were possible. When $\mathscr{R}_0 \geq 1/(1-\theta)$, however, the disease is so infectious relative to vaccine efficacy that eradication cannot be achieved even if all newborns are vaccinated. The government settles for the goal it can achieve, universal vaccination. Equation (22) characterizes the highest price at which all consumers are still willing to purchase, which is associated with

[15] Additional formal results help gauge how substantial the nonmonotonicity in equilibrium variables can be. One can show that the interior maximum of I_m^* as a proportion of its limiting value $\lim_{\mathscr{R}_0 \uparrow \infty} I_m^*$ can be made unboundedly large for suitably chosen parameters. The interior minimum for W_m^* over $\mathscr{R}_0$ can be up to 25% larger than $\lim_{\mathscr{R}_0 \uparrow \infty} W_m^*$.

Table 4
Optimal Subsidy G^{**} as a Function of $\mathscr{R}_0$.

Variable	(a) $\mathscr{R}_0 \in [0,1]$	(b′) $\mathscr{R}_0 \in \left(1, \frac{1}{1-\theta}\right]$	(c′) $\mathscr{R}_0 \in \left(\frac{1}{1-\theta}, \frac{1}{(1-\theta)(1-\tilde{c})}\right]$	(d) $\mathscr{R}_0 \in \left(\frac{1}{(1-\theta)(1-\tilde{c})}, \frac{1}{(1-\theta)^2(1-\tilde{c})}\right]$	(e) $\mathscr{R}_0 \in \left(\frac{1}{(1-\theta)^2(1-\tilde{c})}, \infty\right)$
G_c^{**}	0	c	$c+\theta H\left[\frac{1}{(1-\theta)\mathscr{R}_0}-1\right]$	0	0
G_m^{**}	0	$c+\theta H(\mathscr{R}_0-1)$	$c+\theta H\left[\frac{1}{(1-\theta)^2\mathscr{R}_0}-1\right]$	$c+\theta H\left[\frac{1}{(1-\theta)^2\mathscr{R}_0}-1\right]$	0

Notes: Cases (a), (d), and (e) are identical to those in Tables 2 and 3. Cases (b′) and (c′) differ from (b) and (c) in the previous tables in that the boundary between (b′) and (c′) is $1/(1-\theta)$ rather than $1/(1-\tilde{c})$. The 0 entries in cases (d) and (e) under perfect competition and (e) under monopoly would be negative if taxes were allowed since the government could raise some revenue without impairing universal vaccination in these cases. The model rules out taxes by assumption.

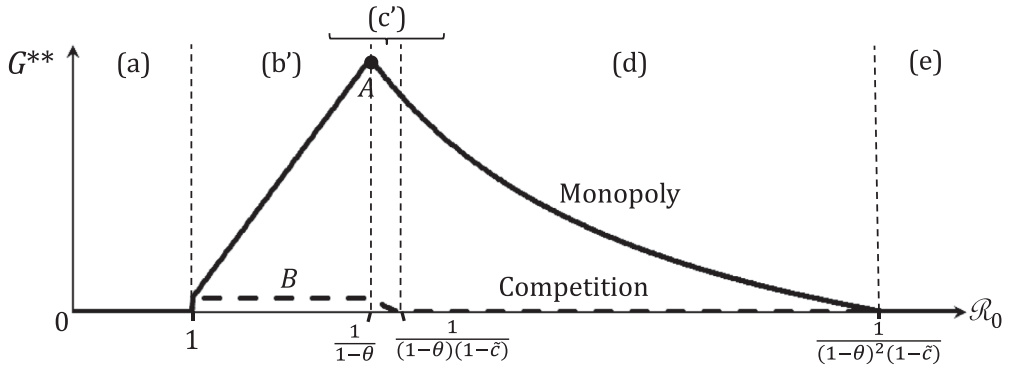

Fig. 2. Graphs of Optimal Subsidy G^{**} as Function of $\mathscr{R}_0$. Notes: Graph of formulas provided in Table 4, illustrated for the specific parameter values indicated in previous figure.

the lowest subsidy required for universal vaccination under the maintained assumption that equilibrium price is weakly decreasing in the subsidy.

With these general principles in hand, we can turn to the analysis of specific market structures. Table 4 provides algebraic expressions for optimal subsidies under perfect competition and monopoly for various values of $\mathscr{R}_0$; Fig. 2 provides graphs. As the figure documents, the optimal subsidy is weakly lower under competition than monopoly. Under both market structures, the optimal subsidy is hump-shaped in $\mathscr{R}_0$, mirroring the hump shape of the marginal externality. The hump shape arises for the same reason, that moderately infectious diseases provide consumers with more leeway to free ride on the vaccination of others.[16]

The next proposition, proved in Appendix A, catalogs relevant observations concerning G^{**}.

Proposition 6. *The following results characterize the optimal subsidy, attaining the first best at minimum government expenditure.*

- $G_m^{**} \geq G_c^{**}$, *with strict inequality for all* $\mathscr{R}_0 \in (1, 1/(1-\theta)^2(1-\tilde{c}))$.
- G_c^{**} *and* G_m^{**} *are weakly increasing in* c *and* θ.
- G_c^{**} *and* G_m^{**} *are hump shaped in* $\mathscr{R}_0 > 0$. *The global maxima of* G_c^{**} *form a plateau of height* c *for* $\mathscr{R}_0$ *in the interval* $(1, 1/(1-\theta))$; G_m^{**} *reaches its global maximum,* $c+\theta^2 H/(1-\theta)$, *at the unique value of* $\mathscr{R}_0 = 1/(1-\theta)$.

If the vaccine were perfectly effective, the optimal subsidy would not be hump shaped but would be weakly increasing over all $\mathscr{R}_0$. The optimal subsidy's hump shape thus hinges on imperfect efficacy. Imperfect efficacy is a maintained assumption in our model and is of practical relevance for currently available vaccines (see Section 6.1). With an imperfectly effective vaccine, vaccine externalities disappear in the limit of an extremely infectious disease as unvaccinated consumers are almost certain to contract the disease from unsuccessfully vaccinated consumers even if the rest of population is vaccinated. No subsidy is needed to obtain the first best with an extremely infections disease and imperfectly effective vaccine.

[16] While G^{**} has the same general hump shape over $\mathscr{R}_0$ as MEX*, the two are not identical for several reasons. In case (e), the first best is obtained without a subsidy under both market structures. Even though all consumers are vaccinated in equilibrium, eliminating the familiar positive externality of vaccination on the unvaccinated, MEX* is still positive, reflecting the externality on the unsuccessfully immunized. Other factors can point in the opposite direction, leading MEX* to exceed G^{**}. MEX* measures the marginal externality at the equilibrium quantity Q^*, but G^{**} must address the externality at the higher quantity Q^{**}. The marginal externality is likely to be higher at the higher quantity since consumers have a lower marginal private benefit, so the externality accounts for a higher fraction of the marginal social benefit. An additional wedge between the optimal subsidy and marginal externality under monopoly is that firms with market power may only incompletely pass through the subsidy.

Saying that the optimal subsidy is weakly lower under perfect competition than monopoly understates the huge gap between global maxima of the curves apparent from Fig. 2. Inspection of Proposition 6 shows that the gap between the global maxima equals $H\theta^2/(1-\theta)$, which grows without bound in the limit of a perfectly effective vaccine.

Appendix B solves for the optimal subsidy under Cournot competition, showing that results in Proposition 6 hold verbatim under Cournot competition for any natural number of competitors n.

5. Vaccines versus drugs

Kremer and Snyder (2015) list a number of legal, technical, behavioral, and operational explanations for why firms may be biased against developing vaccines. The free-rider problem associated with vaccines provides another explanation. The same free-rider problem would not arise for a drug that treat symptoms without inhibiting transmission. Without free riding to reduce demand, the drug would be more lucrative than a similarly effective vaccine, providing stronger development incentives.[17]

This section studies the conditions under which the firm develops the "wrong" product from society's perspective and when the welfare loss from this distortion is greatest. To abstract from the dissipative effect that competition may have on innovation incentives, we focus on the monopoly case throughout the section. The monopoly can produce a vaccine or a drug that is similar in all ways except that it does not reduce disease transmission. We assume both products are costless to produce, i.e., $c = 0$.[18] Let θ denote efficacy for both products. Efficacy for the drug means it eliminates any harm from the symptoms experienced by infected individuals but does not prevent them from transmitting the disease to susceptible individuals. One course of the drug is sufficient to eliminate symptoms for the rest of the consumer's life. If this first course is ineffective for an individual, further courses will be ineffective for that individual as well.

We start by computing monopoly profit and welfare from a drug, respectively Π^*_{md} and W^*_{md}, in steady-state equilibrium. If $\mathscr{R}_0 \leq 1$, the disease naturally dies out in the steady-state, implying $Q^*_{md} = \Pi^*_{md} = 0$. If $\mathscr{R}_0 > 1$, the monopoly can charge $P^*_{md} = \theta H$ for the drug to all newly infected consumers each instant. According to Eq. (3), new infections, βIS, must balance removals from the infected population, $(\alpha+\mu)I$, to maintain $\dot{I} = 0$ in the steady state. Hence, we can compute new infections as $(\alpha+\mu)I(0)$, where the argument added to $I(0)$ indicates that the drug does nothing to curtail infections, as in the vaccine model with no vaccine sales. Equilibrium drug quantity is thus $Q^*_{md} = (\alpha+\mu)I(0) = \mu(1-1/\mathscr{R}_0)$ by (7). Given production is costless,

$$\Pi^*_{md} = P^*_{md}Q^*_{md} = \mu\theta H\left(1-\frac{1}{\mathscr{R}_0}\right). \tag{23}$$

Welfare is

$$W^*_{md} = h[1-I(0)+\theta I(0)] = H\left[\alpha+\mu-\mu(1-\theta)\left(1-\frac{1}{\mathscr{R}_0}\right)\right]. \tag{24}$$

Comparing these expressions against the analogous entries in Table 3 for a vaccine leads to the next proposition, where $\Delta\Pi^*_m = \Pi^*_{md} - \Pi^*_{mv}$ denotes the gap in equilibrium profit between a drug and vaccine, $\Delta W^*_m = W^*_{md} - W^*_{mv}$ the gap in equilibrium welfare between a drug and vaccine, and $\Delta W^{**} = W^{**}_d - W^{**}_v$ the gap in first-best welfare between a drug and vaccine.

Proposition 7. *Suppose $c = 0$. For all $\mathscr{R}_0 > 0$, first-best welfare is weakly lower with a drug than a vaccine ($\Delta W^{**} \leq 0$) and equilibrium profit weakly higher ($\Delta\Pi^*_m \geq 0$). The preceding inequalities are strict if and only if $\mathscr{R}_0 > 1$. The profit advantage from a drug $\Delta\Pi^*_m$ is hump shaped in $\mathscr{R}_0$, approaching its infimum at both extremes, i.e., $\lim_{\mathscr{R}_0\downarrow 0}\Delta\Pi^*_m = \lim_{\mathscr{R}_0\uparrow\infty}\Delta\Pi^*_m = \inf_{\mathscr{R}_0>0} = 0$, and achieving a global maximum $\theta^2H/(1-\theta)$ at the interior value $\mathscr{R}_0 = (1+\theta)^2$. Equilibrium welfare is lower with a drug than a vaccine ($\Delta W^*_m < 0$) if and only if $\mathscr{R}_0 > [\theta/(1-\theta)]^2$.*

The proposition states that the monopoly prefers to develop the drug over the vaccine as long as there is a nontrival market for the products ($\mathscr{R}_0 > 1$). However, if $\mathscr{R}_0 > [\theta/(1-\theta)]^2$, social welfare is higher with a vaccine, implying that the monopoly is biased toward the "wrong" product for sufficiently high $\mathscr{R}_0$. While the drug has the advantage that the monopoly sells the first-best quantity in equilibrium, a drug dose is socially inferior to a vaccine dose because the drug offers no positive externality. Like other variables studied so far that capture the impact of the epidemiological externality on economic outcomes (including the marginal externality and government subsidy), here the magnitude of the monopoly's bias toward the "wrong" product, as quantified by $\Delta\Pi^*_m$, is nonmonotonic in $\mathscr{R}_0$, greatest for some interior value. The externality disappears if the disease is hardly infectious and is swallowed by consumers' private benefit if the disease is infinitely infectious.

[17] If one takes the focus on steady states as literally implying that the discount rate is zero, then there would not be any bias in development decisions. All products generating any positive flow profit would be developed regardless of the size of the up-front development cost. The flow always swamps the up-front cost at a zero discount rate. We are taking the steady-state profit and welfare differentials as approximations of the present discounted value of streams with a positive discount rate.

[18] Besides being convenient, this normalization avoids the issue with positive costs that equating products' unit costs entails different total costs for serving the population. The total population cost for a drug that only needs to be administered to infected consumers is lower than for a vaccine administered to the population in advance of infection.

6. Calibrations

This section provides a series of calibrations to help understand the implications of the analysis for existing vaccines. The calibrations are meant more as illustrations than forecasts. Our present model is too stylized on many fronts to provide accurate forecasts. We abstract from heterogeneity in infectiousness, heterogeneity in costs of prevention among consumers, and mortality effects of disease. Certain parameters are set to convenient limiting values rather than being estimated from data. A host of political-economy considerations lead real-world vaccine markets to depart from our theoretical construct of firms selling directly to individual consumers without third-party funding.

6.1. Parameters

For vaccine efficacy, we take the limiting case of a perfectly effective vaccine, $\theta \uparrow 1$. This limit is interesting on pedagogical grounds because it allows disease eradication to be technologically feasible. This limit is also interesting on practical grounds, reflecting the high efficacy of many existing vaccines.[19] For rescaled cost, $\tilde{c} = c/\theta H$, we take the limiting case of a costless vaccine, $\tilde{c} \downarrow 0$, reflecting the low cost c for existing vaccines, especially in comparison to the potential disease harm H.[20] For the recovery rate, we take the limiting case of no recovery, $\alpha \downarrow 0$, meaning that the person continues to experience harm and can transmit the disease over his or her remaining lifespan, characteristic of diseases such as HIV, syphillis, and malaria. This normalization is not crucial but slightly simplifies one formula.

Imposing these limits and restricting attention to $\mathscr{R}_0 > 1$ considerably simplifies the analysis. Only case (c) remains from Tables 2 and 3 and case (b′) from Table 4. The following proposition is then immediate from inspection of the tables.

Proposition 8. *Suppose $\mathscr{R}_0 > 1$ and consider the limits $\theta \uparrow 1$, $\tilde{c} \downarrow 0$, and $\alpha \downarrow 0$.*

- *Under perfect competition, $P_c^* = \text{MPB}_c^* = \Pi_c^* = I_c^* = G_c^{**} = 0$, $Q_c^* = \mu(1 - 1/\mathscr{R}_0)$, $\text{MEX}_c^* = H$, and $W_c^* = \mu H$.*
- *Under monopoly, $P_m^* = \text{MPB}_m^* = H\left(1 - 1/\sqrt{\mathscr{R}_0}\right)$, $Q_m^* = \mu\left(1 - 1/\sqrt{\mathscr{R}_0}\right)$, $\Pi_m^* = \mu H\left(1 - 1/\sqrt{\mathscr{R}_0}\right)^2$, $I_m^* = \left(1/\sqrt{\mathscr{R}_0}\right) - (1/\mathscr{R}_0)$, $\text{MEX}_m^* = H/\sqrt{\mathscr{R}_0}$, $W_m^* = \mu H\left(1 + 1/\mathscr{R}_0 - 1/\sqrt{\mathscr{R}_0}\right)$, and $G_m^{**} = H(\mathscr{R}_0 - 1)$.*

According to the proposition, under perfect competition, for all $\mathscr{R}_0 > 1$, enough consumers are vaccinated to eradicate the disease, attaining the first best. Price, profit, and the infection rate are all 0. Under monopoly, price, profit, and the share of consumers vaccinated approach 0 in the limit $\mathscr{R}_0 \downarrow 1$, while price approaches 100% of the harm from contracting the disease and quantity approaches 100% share of newborns in the limit $\mathscr{R}_0 \uparrow \infty$. Equilibrium prevalence I_m^* is hump-shaped in $\mathscr{R}_0$, reaching a maximum of $I_m^* = 25\%$ at $\mathscr{R}_0 = 4$. Equilibrium welfare W_m^* is U-shaped, reaching an interior minimum of $W_m^* = (3/4)\mu H$ at $\mathscr{R}_0 = 4$. At this value of $\mathscr{R}_0$, deadweight loss is 25% of first-best welfare, the maximum possible for these parameters.

6.2. HIV calibration

Our first calibration considers a disease, HIV, with a moderate $\mathscr{R}_0$. Table 1 provides a range of estimates for $\mathscr{R}_0$ for HIV. For convenience, we take the round number $\mathscr{R}_0 = 4$ from this range. Incidentally, recall that this is the value of $\mathscr{R}_0$ for which disease prevalence is greatest in equilibrium under monopoly.

Substituting $\mathscr{R}_0 = 4$ into the formulas provided by Proposition 8, the model suggest that a monopoly selling an HIV vaccine to consumers would price it at half the harm from contracting the disease. At this price, half of consumers purchase the vaccine and the other half free ride. Half of the free riders become infected, resulting in an overall HIV prevalence rate of 25%. This is much lower than the 75% prevalence rate that would emerge in the absence of a vaccine for a disease with $\mathscr{R}_0 = 4$. The monopoly captures a quarter of the potential social surplus from a vaccine. Consumers capture one half, and a quarter is deadweight loss.

As noted in the previous subsection, for all $\mathscr{R}_0 > 1$, including $\mathscr{R}_0 = 4$ assumed here, the first best for this essentially costless vaccine involves vaccinating enough people to eradicate the disease in the steady state. The first best is realized under perfect competition. The minimum subsidy that would have to be paid to the monopoly to attain the first best is $G_m^{**} = 3H$. In other words, the monopoly would have to receive a per-course subsidy of at least three times the lifetime harm experienced from certainly contracting HIV, easily amounting to many thousands of dollars per course. The model abstracts from any distortion involved in raising government funds. With any deadweight loss of taxation, such enormous subsidies would likely be prohibitively expensive, forcing governments to use other instruments, such as bulk purchases, to attain the first-best level of vaccination or give up on reaching the first best.

The situation is brighter if the HIV vaccine is competitively supplied since, as mentioned, the first best is obtained in the perfectly competitive equilibrium without a subsidy. Unfortunately, it may be unrealistic to suppose that competition would

[19] For example, U.S. Centers for Disease Control and Prevention, 2020 report efficacies from the recommended vaccine courses of 95% for hepatitis B and tetanus, 97% for measles and shingles, 98% for pertussis, and 99% for polio.

[20] For example, Appendix A3 of Goodkin-Gold et al. (2021) estimates that $\tilde{c} = 2.13 \times 10^{-5}$ in the case of Covid-19.

emerge even in the long run for high-tech product like a potential HIV vaccine that is extremely difficult for a generic competitor to reverse engineer.[21]

6.3. Measles calibration

As a contrast to the HIV calibration, we next calibrate the market for a vaccine for measles, a disease with a much higher value of $\mathscr{R}_0$ than HIV. For convenience, take the round number $\mathscr{R}_0 = 16$ from the range offered by Table 1. The model suggests that a monopoly selling a measles vaccine to consumers would price it at 75% of the harm from contracting the disease. Of consumers, 75% purchase the vaccine, and 25% free ride. A substantial majority, 75%, of free riders contract the disease, resulting in an overall measles prevalence rate of about 19%. This is less than the 25% prevalence calibrated for HIV—as expected given the prevalence rate is maximized at the value of $\mathscr{R}_0$ used in the HIV calibration.

The minimum subsidy required to attain the first best under monopoly is $G_m^{**} = 15H$. In other words, the monopoly would have to receive a per-course subsidy of at least 15 times the lifetime harm experienced from certainly contracting measles. This enormous expense highlights even more strongly than the HIV calibration that the optimal subsidy, while providing a useful theoretical benchmark, would not be a realistic policy alternative in practice for diseases with estimated values of $\mathscr{R}_0$ toward the higher end.

6.4. SARS calibration

Just as the measles calibration provided a useful contrast to HIV, the next calibration for SARS provides a useful contrast but in the other direction. The value $\mathscr{R}_0 = 1.1$ for SARS in Table 1 is considerably lower than for the other diseases, almost lying on the boundary ($\mathscr{R}_0 = 1$) at which the disease disappears in the steady state without a vaccine. The model suggests that a monopoly selling a SARS vaccine to consumers would price it at less than 5% of the harm from contracting the disease. Less than 5% of consumers purchase the vaccine. Almost all consumers free ride. They have the luxury to do so since, despite the low immunization rate, the low infectiousness of SARS keeps equilibrium prevalence low in the calibration at about 4%.

7. Conclusion

Our companion paper (Goodkin-Gold et al., 2021) analyzed a variant of the SIR model tailored to epidemic diseases peaking in a relatively short period against which a concentrated vaccine campaign is waged. This paper adopted the same SIR framework but tailored it to endemic diseases that persist in the population for generations, requiring continuous vaccination. The model may also apply to epidemic diseases if, following Gans (2020), endogenous social distancing is envisaged as flattening epidemic peaks, leading to an extended plateau period during which the infection rate is constant. The analyses of the epidemic and endemic cases provide a more complete picture of vaccine markets for different diseases.

Both papers find that the marginal externality, the optimal subsidy, and the bias against vaccines toward drugs are hump shaped in the infectiousness of the disease as measured by $\mathscr{R}_0$, vanishing for either extremely low or extremely high values. For extremely low $\mathscr{R}_0$, the disease dies out in the steady state; for extremely high $\mathscr{R}_0$, an unvaccinated consumer is almost certain to contract the disease even if everyone else is vaccinated because some of the vaccinated will remain vulnerable due to the vaccine's imperfect efficacy. Both papers also find that the vaccine market performs much worse under monopoly than under perfect competition. The distortion due to the monopoly's market power compounds the free-rider problem.

In this paper, we managed to locate a specific value $\mathscr{R}_0 = 1/(1-\theta)$ at which optimal subsidies are highest under both market structures and more generally under Cournot competition for any natural number of competitors n. This value of $\mathscr{R}_0$ marks the most infectious disease that can be eradicated by vaccinating less than the entire population. Under perfect competition, eradication can only be approached if the entire vaccine cost c is covered by the subsidy since willingness to pay is driven to zero as the disease nears eradication. Under monopoly, an enormous subsidy is required to approach eradication. Monopolies are not in the business of giving away their products; only an enormous subsidy can overcome the monopoly's incentive to restrict output and raise willingness to pay. The subsidy approaches infinity in the limit of a perfectly effective vaccine. The alternative policy of bulk purchases negotiated by the government on behalf of consumers (as done for Covid-19 vaccines as well as the U.S. Vaccines for Children program) may be a better practical alternative.

Kremer and Snyder (2015, 2018) 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. That work suggested that incentives to develop an HIV vaccine may be sufficiently distorted to call for policy attention. That call was reinforced in this paper's HIV calibration, finding that the $\mathscr{R}_0$ estimated for HIV is precisely the value at which deadweight loss and optimal subsidies are maximized given the other parameters.

[21] Certain other health interventions apart from vaccines may be appropriately modeled as being competitively supplied. Consider the use of adult male circumcision as an HIV preventive. Far from universal adoption predicted by the model for a costless, perfectly effective intervention, in the meta-analysis by Kennedy et al. (2020), adult circumcision rates in control samples were negligible. Thus the assumptions of zero cost and perfect efficacy are far-fetched for circumcision and would need to be relaxed.

Our model of the vaccine market is somewhat counterfactual in positing direct-to-consumer vaccine sales. With severe market failures and lives at stake, sensible officials rarely leave vaccine markets alone, so rarely is the unvarnished market observed in practice. Our results on market equilibrium are counterfactual for such markets, but they are a necessary step toward determining the optimal government intervention.

Our focus on subsidies and bulk purchases has left unexamined a range of other policies that might improve the performance of the vaccine market. That the typical welfare results—higher under perfect competition than monopoly, higher under Cournot competition with more competitors—continue to hold in the present setting hints at a possible benefits of using antitrust or compulsory technology licensing to foster competition in the vaccine market. A full analysis of such policies is beyond the scope of this paper. A full analysis would have to account for the impact of the policies on innovation (Finkelstein, 2004) and recognize the difficulty in compelling transfer of a technology as complex as vaccine manufacture, rivaling patents in the strength of intellectual-property protection according to some observers (Eccleston-Turner, 2016).

Our SIR epidemiological model—although standard—has limitations. We make a number of simplifying assumptions including a nonfatal disease and a constant hazard of death, and we abstract from the network structure of transmission. This simplicity allows us to derive closed-form solutions and provide intuitive conditions for our findings, but the model is well short of the complexity required for forecasting purposes, exhibited in such work as Eichenbaum et al. (2020) and Atkeson et al. (2020, 2021). The logic behind the hump shape of the marginal externality and optimal subsidy in $\mathscr{R}_0$ is sufficiently compelling that we expect it to hold in these and the next generation of epidemiological models. We also focus only on steady states rather than undertaking the more complex computation of present discounted values along transition paths, which awaits future work. 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.

Appendix A. Proofs

This appendix supplies proofs not included in the text.

Proof of Proposition 1

Suppose $I^* = 0$ and $Q^* > 0$. We will show $P^* = 0$. Substituting $I^* = 0$ into (7) and rearranging yields $Q^* \geq (\mu/\theta)(1 - 1/\mathscr{R}_0)$. This inequality together with $Q^* > 0$ implies $Q^* \geq \mathscr{Q}_0$ by (8), implying $\text{MPB}^* = 0$ by (11). Since $\text{MPB}(Q)$ is weakly decreasing in Q, $\text{MPB}(\mu) \leq \text{MPB}^* = 0$. Then $Q^* = D(P^*)$ cannot be determined by the third branch of Eq. (13); if it were, $P^* < \text{MPB}(\mu) = 0$, violating the nonnegativity of prices. Since $Q^* > 0$, Q^* cannot be determined by the first branch of (13) either. Therefore,

$$Q^* = d(P^*) = \frac{\mu}{\theta}\left[1 - \frac{1}{(1 - P/\theta H)\mathscr{R}_0}\right]. \tag{A.1}$$

Since $0 = \text{MPB}^* = \text{MPB}(Q^*)$, we have

$$Q^* \geq \mathscr{Q}_0 \geq \frac{\mu}{\theta}\left(1 - \frac{1}{\mathscr{R}_0}\right), \tag{A.2}$$

where the first inequality follows from Eq. (11) and the second from (8). Combining (A.1) with (A.2) and rearranging yields $P^* \leq 0$, implying $P^* = 0$ by the nonnegativity of prices. ■

Proof of Proposition 2

The proof for variables P_c^*, Π_c^*, and MSB_c^*, which are constant over relevant intervals of $\mathscr{R}_0$, are obvious from Table 2.

The reader can verify that Q_c^* is continuous at the boundaries between cases in Table 2. The table entries for Q_c^* are constant for all cases except (c). In this nonempty case, Q_c^* is strictly increasing in $\mathscr{R}_0$. Variables MPB_c^* and W_c^* are analyzed similarly.

The reader can verify that I_c^* is continuous at the boundary between cases in the table and further verify that I_c^* is constant in $\mathscr{R}_0$ in case (a), increasing in case (b), decreasing in case (c), and increasing in cases (d) and (e). A local optimum is thus attained at the boundary between cases (b) and (c). Substituting $\mathscr{R}_0 = 1/(1 - \tilde{c})$ in the case (b) table entry yields $I_c^* = \tilde{c}\mu/(\alpha + \mu)$. The only other candidate for a global maximum is

$$\lim_{\mathscr{R}_0 \uparrow \infty} I_c^* = \frac{(1 - \theta)\mu}{\alpha + \mu}. \tag{A.3}$$

This is greater than the local optimum on the boundary between cases (b) and (c) if and only if $\theta < 1 - \tilde{c}$.

The reader can verify that MEX_c^* is continous at the case (b)–(c) boundary and at the case (c)–(d) boundary and further verify that MEX_c^* is decreasing or constant in $\mathscr{R}_0$ in cases (b)–(e). Hence, MEX_c^* approaches its supremum $\lim_{\mathscr{R}_0 \downarrow 1} \text{MEX}_c^* = 1$ at the boundary between cases (a) and (b). ■

Proof of Proposition 3

The proof for MSB_m^*, which is constant within relevant intervals of $\mathscr{R}_0$, is obvious from Table 3. We omit the comparative-static analysis of Q_m^* as it is similar to that for Q_c^* in the previous proof. We also omit the comparative-static analysis of MPB_m^* and Π_m^*, which are similar to that for Q_m^*. We also omit the comparative-static analysis of MEX_m^*, which is similar to that for MEX_c^* in the previous proof.

We next turn to deriving results for P_m^*. As noted in the text, P_m^* is not unique in cases (a) and (b) in which $Q_m^* = 0$; any price in an interval below c to infinity is consistent with zero output. As stated, we set $P_m^* = c$ by definition in those cases. Thus defined, one can then proceed to analyze the comparative statics of price in a similar way to variables in the preceding paragraph.

We next turn to deriving results for I_m^*. The reader can verify that I_m^* is continuous at the boundary between cases in Table 3 and further verify that I_m^* is constant in $\mathscr{R}_0$ in case (a), increasing in case (b), and increasing in case (e). This leaves cases (c) and (d). Differentiating the table entry for those cases,

$$\frac{\partial I_m^*}{\partial \mathscr{R}_0} = \frac{\mu}{(\alpha+\mu)\mathscr{R}_0^2}\left(1 - \frac{1}{2}\sqrt{\frac{\mathscr{R}_0}{1-\tilde{c}}}\right). \tag{A.4}$$

Eq. (A.4) is decreasing in $\mathscr{R}_0$. It is negative for all $\mathscr{R}_0$ in cases (c) and (d) if it is nonpositive at the lower boundary of case (c). Evaluating (A.4) at this lower boundary $\mathscr{R}_0 = 1/(1-\tilde{c})$, we see it is nonpositive if and only if $\tilde{c} \geq 1/2$. But then I_m^* is nonmonotonic, for it is increasing in case (b) and decreasing in cases (c) and (d). It reaches a local maximum of $\tilde{c}\mu/(\alpha+\mu)$ at the boundary between cases (b) and (c). This weakly exceeds $\lim_{\mathscr{R}_0\uparrow\infty} I_m^* = (1-\theta)\mu/(\alpha+\mu)$ if $\theta \geq 1-\tilde{c}$, in which case the local maximum is a global maximum.

Assume $\tilde{c} < 1/2$. Then (A.4) is nonnegative for all $\mathscr{R}_0$ in cases (c) and (d) if and only if (A.4) is nonnegative at the upper boundary of case (d). Evaluating (A.4) at this upper boundary $\mathscr{R}_0 = 1/(1-\theta)^2(1-\tilde{c})$, we see it is nonnegative if and only if $\theta \leq (1-2\tilde{c})/2(1-\tilde{c})$. Under these conditions, I_m^* is nondecreasing in $\mathscr{R}_0$ for all $\mathscr{R}_0 > 0$.

For the remaining parameters, I_m^* reaches a local maximum in the interior of cases (c) and (d). Setting (A.4) to 0 and solving implies that the local maximum is at $\mathscr{R}_0 = 4(1-\tilde{c})$. This local maximum equals $\mu/4(\alpha+\mu)(1-\tilde{c})$, which exceeds $\lim_{\mathscr{R}_0\uparrow\infty} I_m^* = (1-\theta)\mu/(\alpha+\mu)$ if $\theta \geq (3-4\tilde{c})/4(1-\tilde{c})$, in which case the local maximum is a global maximum.

Summarizing the results for I_m^*, we have shown that I_m^* is weakly decreasing in $\mathscr{R}_0$ if and only if $\tilde{c}$ and θ are sufficiently low, in particular, $\tilde{c} < 1/2$ and $\theta \leq (1-2\tilde{c})/2(1-\tilde{c})$. Otherwise, I_m^* is nonmonotonic, achieving a local maximum, which is a global maximum for θ above the specified thresholds.

We finally turn to deriving results for W_m^*. The reader can verify that W_m^* is continuous at the boundary between cases in Table 3 and further verify that W_m^* is constant in $\mathscr{R}_0$ in case (a), decreasing in case (b), and decreasing in case (e). This leaves cases (c) and (d). Differentiating the table entry for those cases,

$$\frac{\partial W_m^*}{\partial \mathscr{R}_0} = \frac{\mu H}{\mathscr{R}_0^2}\left(\frac{1}{2}\sqrt{(1-\tilde{c})\mathscr{R}_0} - 1\right). \tag{A.5}$$

The sign of (A.5) depends on the factor in parentheses, which is increasing in $\mathscr{R}_0$. It is nonpositive for all $\mathscr{R}_0$ in cases (c) and (d) if and only if it is nonpositive at the upper boundary of case (d). Evaluating (A.4) at this upper boundary $\mathscr{R}_0 = 1/(1-\theta)^2(1-\tilde{c})$, we see it is nonpositive if and only if $\theta \leq 1/2$. Otherwise, W_m^* reaches a local minimum in the interior of cases (c) and (d). Setting (A.5) to 0 and solving yields a local minimum at $\mathscr{R}_0 = 4/(1-\tilde{c})$. This local minimum equals $H[\alpha + (3/4)(1-\tilde{c})\mu]$, which weakly exceeds $\lim_{\mathscr{R}_0\uparrow\infty} W_m^* = H[\alpha + \mu\theta(1-\tilde{c})]$ if $\theta \geq 3/4$, in which case the local minimum is a global minimum. ■

Proof of Proposition 4

Suppose $\tilde{c} \geq 1/2$. Then I_m^* achieves an interior maximum regardless of θ as shown in the previous proof. So suppose $\tilde{c} < 1/2$. Then the previous proof shows that I_m^* achieves an interior maximum if $\tilde{c} > (1-2\theta)/2(1-\theta)$, or rearranging, $\theta > (1-2\tilde{c})/2(1-\tilde{c})$. The right-hand side of this last inequality is decreasing in $\tilde{c}$. Substituting $\tilde{c} = 0$ gives the sufficient condition $\theta > 1/2$ for I_m^* to achieve an interior maximum. Proposition 3 states that $\theta > 1/4$ is sufficient for W_m^* to achieve an interior minimum. Thus, $\theta > 1/2$ is sufficient for both I_m^* and W_m^* to achieve interior optima.

Suppose $\tilde{c} \geq 1/2$. The previous proof showed that I_m^* achieves a interior global maximum if $\theta \geq 1-\tilde{c} \geq 1/2$. Suppose instead that $\tilde{c} < 1/2$. The previous proof showed that I_m^* achieves an interior global optimum if $\theta \geq (3-4\tilde{c})/4(1-\tilde{c})$. The right-hand side of this last inequality is decreasing in $\tilde{c}$. Substituting $\tilde{c} = 0$ gives the sufficient condition $\theta > 3/4$ for I_m^* to achieve an interior global maximum in this case. Combining the sufficient condition derived supposing $\tilde{c} \geq 1/2$ with that derived supposing $\tilde{c} < 1/2, \theta \geq 3/4$ is a sufficient condition for I_m^* to attain an interior global maximum. As shown in the previous proof, $\theta \geq 3/4$ is also sufficient for W_m^* to achieve an interior global minimum. ■

Proof of Proposition 5

Suppose $\mathscr{R}_0 \leq 1$. Then $Q^{**} = \min(\mathscr{Q}_0, \mu) = \min(0, \mu) = 0$, where the first equality follows from Eq. (20), the second from substituting $\mathscr{R}_0 \leq 1$ into (8), and the third from $\mu > 0$. No subsidy is needed to generate zero quantity, implying $G^{**} = 0$.

Suppose $1 < \mathscr{R}_0 < 1/(1-\theta)$. Substituting $1 < \mathscr{R}_0$ into (8) yields $\mathscr{Q}_0 > 0$, implying $Q^{**} = \min(\mathscr{Q}_0, \mu) > 0$. Substituting $\mathscr{R}_0 < 1/(1-\theta)$ into (8) yields $\mathscr{Q}_0 < \mu$, implying $Q^{**} = \min(\mathscr{Q}_0, \mu) = \mathscr{Q}_0$, in turn implying $I^{**} = 0$ by (7). By Proposition 1, $I^{**} = 0$ and $Q^{**} > 0$ imply $P^{**} = 0$.

Suppose $\mathscr{R}_0 \geq 1/(1-\theta)$. Then $\mathscr{Q}_0 \geq \mu$ by (8), implying $Q^{**} = \min(\mathscr{Q}_0, \mu) = \mu$. To find the lowest subsidy delivering $Q^{**} = \mu$, we need to find the highest P^{**} satisfying $D(P^{**}) = \mu$ since the equilibrium price is weakly decreasing in the subsidy.

Consider a price P' such that $P' > \text{MPB}(\mu)$. Then $D(P') \leq d(P')$ since only the first two branches of (13) are relevant in the computation of $D(P')$. Combining the inequality

$$P' > \text{MPB}(\mu) = \frac{\theta H}{1-\theta}\left(1 - \theta - \frac{1}{\mathscr{R}_0}\right) \tag{A.6}$$

with (12) yields $d(P') < \mu$. Thus, $D(P') \leq d(P') < \mu = Q^{**} = D(P^{**})$, implying $P^{**} \neq P'$. But since P' was an arbitrary price greater than $\text{MPB}(\mu)$, we have $P^{**} \leq \text{MPB}(\mu)$. The highest price satisfying $P^{**} \leq \text{MPB}(\mu)$ is $P^{**} = \text{MPB}(\mu)$. To verify that this price yields the desired quantity, we have $\mu = d(\text{MPB}(\mu)) = D(\text{MPB}(\mu)) = D(P^{**})$, where the first equality follows from substituting from (A.6) into (12) and the second equality from the fact that only the middle branch is relevant in (13) at a price of $\text{MPB}(\mu)$.

The previous paragraph shows $d(P^{**}) = d(\text{MPB}(\mu)) = \mu$, implying $P^{**} = d^{-1}(\mu)$. Inverting (12) yields Eq. (22). ■

Proof of Proposition 6

We first derive the expressions for G_c^{**} and G_m^{**} appearing in Table 4 from which many of the results in the proposition are gleaned. Under perfect competition, the firms pass the subsidy directly to consumers, implying $P^{**} = c - G_c^{**}$ and thus $G_c^{**} = c - P^{**}$. Substituting the relevant values of P^{**} from Proposition 5 yields $G_c^{**} = 0$ if $\mathscr{R}_0 \leq 1$, $G_c^{**} = c$ if $\mathscr{R}_0 \in (1, 1/(1-\theta))$, and

$$G_c^{**} = c - \theta H\left[1 - \frac{1}{(1-\theta)\mathscr{R}_0}\right] \tag{A.7}$$

if $\mathscr{R}_0 \geq 1/(1-\theta)$.

For an imperfectly effective vaccine, (A.7) becomes negative for sufficiently large $\mathscr{R}_0$. Had we not ruled out negative subsidies by assumption, these negative values of (A.7) would indeed constitute G_c^{**}. Given the nonnegativity constraint on subsidies, we have $G_c^{**} = 0$ for $\mathscr{R}_0 \geq 1/(1-\theta)(1-\tilde{c})$. It is no coincidence that this is same threshold for the perfectly competitive equilibrium to obtain the first best in the absence of a subsidy. No subsidy is needed if equilibrium generates the first best without one. For $\mathscr{R}_0$ strictly above this threshold, the government would like to tax vaccines since some revenue can be raised without impairing universal vaccination.

Next turn to computing the optimal subsidy under monopoly. By Proposition 5, $G_m^{**} = 0$ if $\mathscr{R}_0 \leq 1$. If $\mathscr{R}_0 \in (1, 1/(1-\theta))$, Proposition 5 implies $Q^{**} = \mathscr{Q}_0$. Setting the quantity in (21) derived from the monopoly's first-order condition equal to $\mathscr{Q}_0$, substituting the effective marginal cost $c - G$ under a subsidy for c in the formula, and rearranging yields $G_m^{**} = c + \theta H(\mathscr{R}_0 - 1)$. If $\mathscr{R}_0 \geq 1/(1-\theta)$, calculations are similar except that the monopoly quantity in (21) needs to equated with the relevant first-best quantity in this case, $Q^{**} = \mu$, yielding

$$G_m^{**} = \max\left\{0, c + \theta H\left[\frac{1}{(1-\theta)^2\mathscr{R}_0} - 1\right]\right\}, \tag{A.8}$$

where the max operator has been added to reflect the nonnegativity constraint on subsidies. The nonnegativity constraint binds in case (e)—not coincidentally the case in which the first-best would be obtained in monopoly equilibrium without a subsidy.

It remains to prove the second bullet point regarding comparative statics of G^{**} in c and θ. Gleaning the remaining results from Table 4 is relatively straightforward. In case (a) of the table, $G_c^{**} = G_m^{**} = 0$, in which case the comparative statics hold trivially. The remaining cases can be combined in a single expression for each market structure,

$$G_c^{**} = \min\left\{c, \max\left\{0, c + \theta H\left[\frac{1}{(1-\theta)\mathscr{R}_0} - 1\right]\right\}\right\} \tag{A.9}$$

$$G_m^{**} = \min\left\{c + \mathscr{R}_0 - 1, \max\left\{0, c + \theta H\left[\frac{1}{(1-\theta)^2\mathscr{R}_0} - 1\right]\right\}\right\}, \tag{A.10}$$

which are obviously weakly increasing in c and θ. ■

Proof of Proposition 7

Suppose $\mathscr{R}_0 \leq 1$. Then $Q_{md}^* = Q_{mv}^* = 0$, implying $\Pi_{md}^* = \Pi_{mv}^* = 0$ and $W_{md}^* = W_{mv}^* = W_d^{**} = W_d^{**} = (\alpha + \mu)H$, implying $\Delta\Pi_m^* = \Delta W_m^* = \Delta W^{**} = 0$. The fact that $\Delta\Pi_m^* = 0$ for all $\mathscr{R}_0 < 1$ implies $\lim_{\mathscr{R}_0 \downarrow 0} \Delta\Pi_m^* = 0$.

Suppose $\mathscr{R}_0 > 1$. The assumption $c = 0$ implies $\tilde{c} = 0$, leaving two cases in Table 2: (c)–(d) and (e). In case (c)–(d), defined by

$$\mathscr{R}_0 \in \left(1, \frac{1}{(1-\theta)^2}\right), \tag{A.11}$$

substituting from (23) for Π^*_{md} and from the table entry for Π^*_{mv} yields, after rearranging,

$$\Delta\Pi^*_m = \frac{\sqrt{\mathscr{R}_0}-1}{\theta\mathscr{R}_0}\left[1+\theta-(1-\theta)\sqrt{\mathscr{R}_0}\right]. \tag{A.12}$$

The first factor is positive since $\sqrt{\mathscr{R}_0} > 1$ by (A.11); the second is positive since $\sqrt{\mathscr{R}_0}(1-\theta) < 1$ by (A.11). In case (e), $\Delta\Pi^*_m = \theta/(1-\theta)\mathscr{R}_0$, implying $\lim_{\mathscr{R}_0\uparrow\infty}\Delta\Pi^*_m = 0$. Combined with the results for $\mathscr{R}_0 \le 1$, we have that $\Delta\Pi^*_m \ge 0$ for all $\mathscr{R}_0 > 0$ with strict inequality if and only if $\mathscr{R}_0 > 1$. Further, $\inf_{\mathscr{R}_0>0} = 0$.

To verify the quasiconcavity of $\Delta\Pi^*_m$, $\Delta\Pi^*_m$ is a constant 0 in (a). In case (c)–(d),

$$\frac{\partial\Delta\Pi^*_m}{\partial\mathscr{R}_0} = \frac{H}{\mathscr{R}_0^2}\left(1+\theta-\sqrt{\mathscr{R}_0}\right), \tag{A.13}$$

implying that $\Delta\Pi^*_m$ is first increasing, reaches a critical point at $\mathscr{R}_0 = (1+\theta)^2$, and then is decreasing. One can verify that the critical point is in the interior of case (c)–(d), as $(1+\theta)^2 < 1/(1-\theta)^2$. At the boundary of (e), $\Delta\Pi^*_m$ is continuous and continues to decline throughout case (e), proving $\Delta\Pi^*_m$ is quasiconcave for all $\mathscr{R}_0 > 0$.

Turning to ΔW^*_m, in case (e), substituting from (24) for W^*_{md} and from the relevant table entry for W^*_{mv} yields $\Delta W^*_m = -\theta h/\mathscr{R}_0$, which is negative. In case (c)–(d), substituting the relevant table entry for W^*_{mv} in case (c)–(d) yields

$$\Delta W^*_m = \frac{h}{\mathscr{R}_0}\left[(\theta-1)(\sqrt{\mathscr{R}_0})^2 + \sqrt{\mathscr{R}_0} - \theta\right]. \tag{A.14}$$

The sign is determined by the factor in brackets, a quadratic equation in $\sqrt{\mathscr{R}_0}$, which is negative if $\sqrt{\mathscr{R}_0}$ lies outside the roots 1 and $\theta/(1-\theta)$. Since $\mathscr{R}_0 > 1$, the relevant condition is $\sqrt{\mathscr{R}_0} > \theta/(1-\theta)$, implying $\mathscr{R}_0 > [\theta/(1-\theta)]^2$, the stated condition for $\Delta W^*_m < 0$.

Turning to ΔW^{**}, the first-best quantity is sold in equilibrium with a drug, implying $W^{**}_d = W^*_{md}$. By (20), the first-best vaccine quantity is $Q^{**} = \mathscr{Q}_0 = (\mu/\theta)(1-1/\mathscr{R}_0)$ if $\mathscr{R}_0 \in (1, 1/(1-\theta)]$ and $Q^{**} = \mu$ if $\mathscr{R}_0 > 1/(1-\theta)$. Suppose $\mathscr{R}_0 \in (1, 1/(1-\theta)]$. Then $W^{**}_v = (\alpha+\mu)H$. Substituting this value along with the value of $W^*_d = W^{**}_d$ from (24) and rearranging yields $\Delta W^{**} = W^{**}_d - W^{**}_v = -\mu H(1-\theta)(1-1/\mathscr{R}_0) < 0$. Next, suppose $\mathscr{R}_0 > 1/(1-\theta)$. Then, according to the case (e) entry in Table 3 setting $\tilde{c} = 0, W^{**}_v = H(\alpha+\theta\mu+\mu/\mathscr{R}_0)$. Substituting this value along with the value of $W^*_d = W^{**}_d$ from (24) and rearranging yields $\Delta W^{**} = W^{**}_d - W^{**}_v = -\theta\mu H/\mathscr{R}_0$. Thus, $\Delta W^{**} \le 0$ for all $\mathscr{R}_0 > 0$ with strict inequality for $\mathscr{R}_0 > 1$. ■

Appendix B. Cournot competition

Cases (a) and (b) from Table 2, which involved no sales under perfect competition, will also involve no sales under Cournot since firms mark up marginal costs. Thus the entries in cases (a) and (b) from both Tables 2 and 3 will also apply to Cournot.

For the remainder of this appendix, suppose $\mathscr{R}_0 > 1/(1-\tilde{c})$. Letting q_i denote firm i's output and Q_{-i} is the output of i's rivals, i's profit equals $[P(q_i+Q_{-i})-c]q_i = [\mathrm{MPB}(q_i+Q_{-i})-c]q_i$. 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

$$Q^*_n = \frac{\mu}{\theta}\left(1-\frac{\psi}{\mathscr{R}_0}\right), \tag{B.1}$$

where

$$\psi(n) = \frac{n-1+\sqrt{(n-1)^2+4n(1-\tilde{c})\mathscr{R}_0}}{2n(1-\tilde{c})}. \tag{B.2}$$

While ψ is a function of variables besides n, in particular $\tilde{c}$ and $\mathscr{R}_0$, those are supressed as arguments of ψ for brevity.

Table B.1 records the expression (B.1) for equilibrium quantity in the relevant row (second row is the relevant one for quantity) and column (penultimate column is the relevant one for the parametric case under consideration). The rest of the variables in the penultimate column can be derived by substituting Q^*_n for Q in the relevant formulas provided in the text.

The preceding analysis is valid if

$$\mathscr{R}_0 \le \frac{1+(1-\theta)(n-1)}{n(1-\theta)^2(1-\tilde{c})}. \tag{B.3}$$

Otherwise, $Q^*_n > \mu$ for the Q^*_n in (B.1). 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 $Q^*_n = \mu$. The rest of the equilibrium

Table B.1
Steady-state equilibrium variables under Cournot as functions of $\mathscr{R}_0$.

Variable	(a) $\mathscr{R}_0 \in [0,1]$	(b) $\mathscr{R}_0 \in \left(1, \frac{1}{1-\tilde{c}}\right]$	(c), (d′) $\mathscr{R}_0 \in \left(\frac{1}{1-\tilde{c}}, \frac{1+(1-\theta)(n-1)}{n(1-\theta)^2(1-\tilde{c})}\right]$	(e′) $\mathscr{R}_0 \in \left(\frac{1+(1-\theta)(n-1)}{n(1-\theta)^2(1-\tilde{c})}, \infty\right)$
P_n^*	c	c	$\theta H\left[1-\frac{1}{\psi(n)}\right]$	$\theta H\left[1-\frac{1}{(1-\theta)\mathscr{R}_0}\right]$
Q_n^*	0	0	$\frac{\mu}{\theta}\left[1-\frac{\psi(n)}{\mathscr{R}_0}\right]$	μ
Π_n^*	0	0	$\frac{\mu H}{\theta}\left[1-\tilde{c}-\frac{1}{\psi(n)}\right]\left[1-\frac{\psi(n)}{\mathscr{R}_0}\right]$	$\mu\theta H\left[1-\tilde{c}-\frac{1}{(1-\theta)\mathscr{R}_0}\right]$
I_n^*	0	$\frac{\mu}{\alpha+\mu}\left(1-\frac{1}{\mathscr{R}_0}\right)$	$\frac{\mu}{\alpha+\mu}\left[\frac{\psi(n)-1}{\mathscr{R}_0}\right]$	$\frac{\mu}{\alpha+\mu}\left(1-\theta-\frac{1}{\mathscr{R}_0}\right)$
MPB_n^*	0	$\theta H\left(1-\frac{1}{\mathscr{R}_0}\right)$	$\theta H\left[1-\frac{1}{\psi(n)}\right]$	$\theta H\left[1-\frac{1}{(1-\theta)\mathscr{R}_0}\right]$
MSB_n^*	0	θH	θH	θH
MEX_n^*	0	$\frac{\theta H}{\mathscr{R}_0}$	$\frac{\mu H}{\psi(n)}$	$\frac{\theta H}{(1-\theta)\mathscr{R}_0}$
W_n^*	$H(\alpha+\mu)$	$H\left(\alpha+\frac{\mu}{\mathscr{R}_0}\right)$	$H\left\{\alpha+\mu\left\{(1-\tilde{c})\left[1-\frac{\psi(n)}{\mathscr{R}_0}\right]+\frac{1}{\mathscr{R}_0}\right\}\right\}$	$H\left[\alpha+\mu\theta(1-\tilde{c})+\frac{\mu}{\mathscr{R}_0}\right]$

Notes: The distinction between cases (c) and (d′), relevant for perfect competition in Table 2, is irrelevant for Cournot here. In cases (a) and (b) in which equilibrium output is zero, any price in an interval from below c to infinity is consistent with zero output. As discussed in footnote 11, we adopt the convention that $P_n^* = c$ in those cases.

variables have the same formula as in case (e) of Table 3. Note that the threshold between cases (d) and (e) is different, given by the right-hand side of (B.3). Table B.1 records equilibrium variables under Cournot competition for all values of $\mathscr{R}_0$.

One can show that $\lim_{n\uparrow\infty}\psi(n) = 1/(1-\tilde{c})$ and thus that the preceding expressions for the equilibrium variables collapse to their values under perfect competition given in Table 2. One can also show that $\psi(n) = \sqrt{\mathscr{R}_0}/\sqrt{1-\tilde{c}}$ for $n = 1$, and thus that the preceding expressions for the equilibrium variables collapse to their monopoly values given in Table 3.

We can also show that $\psi'(n) < 0$ for $\mathscr{R}_0 > 1/(1-\tilde{c})$. We have

$$\psi'(n) = \frac{\sqrt{(n-1)^2 + 4n(1-\tilde{c})\mathscr{R}_0} - [2n(1-\tilde{c})\mathscr{R}_0 - (n-1)]}{n^2\sqrt{(n-1)^2 + 4n(1-\tilde{c})\mathscr{R}_0}} \tag{B.4}$$

The denominator of (B.4) is positive. One can show the numerator is negative by showing that the square of the second factor exceeds the radicand in the first factor when $\mathscr{R}_0 > 1/(1-\tilde{c})$. Hence, $\psi'(n) < 0$ for $\mathscr{R}_0 > 1/(1-\tilde{c})$. In view of this fact, inspection of Table B.1 shows that the effects of increasing n that one would expect to hold in a Cournot model do hold for the variables in the column headed (c) and (d′): P_n^* and I_n^* decrease and Q_n^* and W_n^* increase. The one variable for which the comparative-statics result is difficult to gauge by inspection is Π_n^*. It can be shown that industry profit $\Pi(Q)$ is quasiconcave in Q, implying that increases in Q above the industry-profit maximizer Q_m^* reduce $\Pi(Q)$. As $Q_n^* \geq Q_m^*$ and Q_n^* is increasing in n, Π_n^* is decreasing in n in cases (c) and (d′). Outside of cases (c) and (d′), variables do not change with n. Thus, the comparative-statics effects that hold strictly in the one column hold weakly across columns.

Following the steps used in the proof of Proposition 3 for monopoly, substituting the expressions for equilibrium variables in Table B.1 for those in Table 3, one can show that Proposition 3 holds verbatim for Cournot competition for all natural numbers n.

Finally, we solve for the optimal subsidy under Cournot, G_n^{**} using similar arguments used for monopoly in the proof of Proposition 6. For $\mathscr{R}_0 < 1/(1-\theta)$, $G_n^{**} = G_m^{**}$. For $\mathscr{R}_0 \geq 1/(1-\theta)$, steps leading to the expression for G_m^{**} in (A.9) can be used to derive

$$G_n^{**} = \min\left\{c + \mathscr{R}_0 - 1, \max\left\{0, c + \theta H\left[\frac{n-(n-1)\theta}{n(1-\theta)^2\mathscr{R}_0} - 1\right]\right\}\right\}. \tag{B.5}$$

We thus have that G_n^{**} reaches its global maximum, $c + \theta^2 H/(1-\theta)$, at $\mathscr{R}_0 = 1/(1-\theta)$, the same as G_m^{**}.

Appendix C. Consumer heterogeneity

This section extends the model to heterogeneous consumers, showing that key nonmonotonicities continue to hold. We also explore whether consumer heterogeneity introduces bias into naïve estimates of the population mean values of key variables such as the marginal externality and, if so, how to correct for it.

Heterogeneity in Benefits

For concreteness, assume consumers, indexed by i, differ in disease harm, H_i. Similar analysis applies if consumers experience different efficacies θ_i or have different lifespans.[22] Denote the probability density function (pdf) by $f(H_i)$, the cumulative distribution function (cdf) by $F(H_i)$, and the complementary cdf by $\bar{F}(H_i) = 1 - F(H_i)$, and the expected value by $E(H_i) = \int_0^\infty H_i f(H_i) dH_i$. Assume H_i has full support on $(0, \infty)$. Assume further that the population distribution of H_i is common knowledge but the specific realization of H_i is consumer i's private information.[23]

With consumer heterogeneity, consumer i's marginal private benefit becomes $\text{MPB}_i(Q) = \theta H_i \Phi(Q)$. Incorporating heterogeneity in the normative measures requires additional work to keep track of the high-value consumers who end up purchasing. Aggregate health equals

$$\text{AH}(Q) = \mu\left\{\left[1 - \Phi(Q)\right]\int_0^{\hat{H}} H_i f(H_i) dH_i + \left\{\theta + (1-\theta)[1 - \Phi(Q)]\right\}\int_{\hat{H}}^\infty H_i f(H_i) dH_i\right\}. \tag{C.1}$$

The first integral reflects the expected health experienced by those whose harm is below the threshold $\hat{H}$ for purchase. With no vaccine to protect them, consumer i in this group obtains H_i with probability $1 - \Phi(Q)$. The second integral reflects the expected health experienced by purchasers. Purchaser i obtains H_i if he or she is successfully immunized (probability θ) or is not successfully immunized (probability $1 - \theta$) and thus is susceptible but still fails to contract the disease conditional on being susceptible (probability $1 - \Phi(Q)$). The initial factor μ scales the per-consumer surplus in braces by the population of potential consumers. With no vaccine, (C.1) reduces to

$$\text{AH}(0) = \mu\left[1 - \Phi(0)\right]\int_0^\infty H_i f(H_i) dH_i. \tag{C.2}$$

Subtracting (C.2) from (C.1) to compute social benefit from a vaccine $\text{SB}(Q) = \text{AH}(Q) - \text{AH}(0)$ yields, after rearranging,

$$\text{SB}(Q) = \mu\left\{\left[\Phi(0) - \Phi(Q)\right]E(H_i) + \theta\Phi(Q)\int_{\hat{H}}^\infty H_i f(H_i) dH_i\right\}. \tag{C.3}$$

Differentiating (C.3) yields

$$\text{MSB}(Q) = \mu\left\{-\frac{\partial\Phi(Q)}{\partial Q}\left[E(H_i) - \theta\int_{\hat{H}}^\infty H_i f(H_i) dH_i\right] - \theta\Phi(Q)\hat{H}f(\hat{H})\frac{\partial\hat{H}}{\partial Q}\right\}. \tag{C.4}$$

To compute $\partial\hat{H}/\partial Q$, note threshold consumer type $\hat{H}$ is given as an implicit function of Q by $Q = \mu\bar{F}(\hat{H})$. Totally differentiating this identity with respect to Q and rearranging yields $\partial\hat{H}/\partial Q = -1/\mu f(\hat{H})$. After substituting this derivative into (C.4), we see that the last term equals $\theta\hat{H}\Phi(Q)$. This is the private benefit of the threshold consumer, equal to MPB^* when evaluated at the equilibrium Q^*. Subtracting to compute $\text{MEX}^* = \text{MSB}^* - \text{MPB}^*$ leaves just the first term of (C.4), as stated in the following lemma.

Lemma 1. *In the model with heterogeneity in consumer harm H_i, the marginal externality is*

$$\text{MEX}^* = -\frac{\partial\Phi(Q^*)}{\partial Q}\left[E(H_i) - \theta\int_{\hat{H}(Q^*)}^\infty H_i f(H_i) dH_i\right]\mu. \tag{C.5}$$

Intuitively, Lemma 1 says that the marginal externality is proportional to $-\partial\Phi(Q^*)/\partial Q$, the decline in the equilibrium probability of infection for an unvaccinated individual when one additional susceptible is vaccinated. The proof of the next proposition shows that that leading factor approaches 0 as $\mathscr{R}_0 \downarrow 0$ in both long- and short-run analyses since a noninfectious disease presents no danger of infection in either analysis. The factor also approaches 0 as $\mathscr{R}_0 \uparrow \infty$ in both analyses since the individual will almost certainly contract the infinitely infectious disease in any event—from someone who was vaccinated but for whom the vaccine was ineffective if no one else. The remaining factors are obviously positive and finite for all $\mathscr{R}_0$. Thus, MEX^* approaches 0 for extreme values of $\mathscr{R}_0$, implying it is nonmonotonic in $\mathscr{R}_0$, as the following proposition states.

Proposition 9. *In the model with heterogeneity in consumer harm H_i, under both perfect competition and monopoly, MEX^* approaches 0 in both extremes $\mathscr{R}_0 \downarrow 0$ and $\mathscr{R}_0 \uparrow \infty$ and achieves an interior global maximum in $\mathscr{R}_0$.*

Proof. It remains to show that the limit of $\partial\Phi(Q^*)/\partial Q$ as $\mathscr{R}_0$ approaches either extreme equals 0. One can use direct computation to verify that the expression for $\Phi(Q)$ provided in Eq. (10) for homogeneous consumers continues to apply in the model of heterogeneous consumers here. For $\mathscr{R}_0 \leq 1, \mathscr{Q}_0 = 0$, implying $\Phi(Q) = 0$ for all Q by (10), implying $\partial\Phi(Q)/\partial Q = 0$ for all Q, implying $\lim_{\mathscr{R}_0\downarrow 0}\partial\Phi(Q^*)/\partial Q = 0$ for both $Q^* = Q_c^*$ and $Q^* = Q_m^*$.

[22] We conjecture that the analysis is also similar if consumers contract the disease at different rates, but modeling heterogeneity in that dimension requires delicacy to avoid changing the epidemiological process.

[23] The model requires consumers to be aware of their heterogeneity, for example, differences in income leading to different willingnesses to pay to avoid harm, or a family history of disease. Undiagnosed conditions that lead harm to vary but are unknown to the consumer are better accommodated in the homogeneous-harm model.

For $\mathscr{R}_0 \uparrow \infty$, $\mathscr{Q}_0 = \mu/\theta > \mu$ for $\theta < 1$, implying the second branch of (10) is the relevant one. Differentiating,

$$\frac{\partial \Phi(Q)}{\partial Q} = \frac{\theta}{\mu \mathscr{R}_0 (1 - \theta Q/\mu)^2}. \tag{C.6}$$

Since $Q \in [0, \mu]$,

$$\frac{\theta}{\mu \mathscr{R}_0} \leq \frac{\partial \Phi(Q)}{\partial Q} \leq \frac{\theta}{\mu \mathscr{R}_0 (1-\theta)^2}. \tag{C.7}$$

Taking limits,

$$\lim_{\mathscr{R}_0 \uparrow \infty} \frac{\theta}{\mu \mathscr{R}_0} \leq \lim_{\mathscr{R}_0 \uparrow \infty} \frac{\partial \Phi(Q)}{\partial Q} \leq \lim_{\mathscr{R}_0 \uparrow \infty} \frac{\theta}{\mu \mathscr{R}_0 (1-\theta)^2}. \tag{C.8}$$

The limits on the far left and far right-hand sides of (C.8) equal 0, implying $\lim_{\mathscr{R}_0 \uparrow \infty} \partial \Phi(Q)/\partial Q = 0$ for all $Q \in [0, \mu]$, including $Q = Q_c^*$ and $Q = Q_m^*$. ■

Aggregate estimation on heterogeneous groups

This section discusses conditions under which a naïve estimate that does not take into account population heterogeneity can recover the population means of the marginal externality and other equilibrium variables.

To derive formal results, consider a simple setting of a population consisting of epidemiologically distinct subgroups indexed by $g \in \Gamma$. Assume the subgroups sharing all parameters in common except for subgroup size, denoted w_g, with $\sum_{g \in G} w_g = 1$, and disease transmissibility β_g, generating different values of the basic reproductive number, $\mathscr{R}_{0g}$ according to Eq. (6). A researcher, unaware of the subgroup heterogeneity, seeks to estimate $\mathscr{R}_0$ and MEX using aggregate data. Assume no vaccine is yet available—skirting a potentially awkward situation in which the expert knows less about the disease than ordinary consumers, who are required to by our economic model to have rational expectations about the disease's epidemiology when making their consumption decisions. The researcher seeks an estimate of the externality associated with a first dose of possible vaccine to be developed.

Assume the researcher estimates $\mathscr{R}_0$—in what would be a model-consistent way absent consumer heterogeneity—by substituting an observation of aggregate, steady-state disease prevalence I and known values of parameters besides $\mathscr{R}_0$ into Eq. (7) and solving for $\mathscr{R}_0$. The researcher then translates the estimate of $\mathscr{R}_0$ into an estimate of MEX using the relevant formula from Tables 2 or 3.

To isolate the effect of consumer heterogeneity, the original epidemiological model needs to be normalized to avoid the outcome that separation itself—even of a homogeneous population—has real effects on $\mathscr{R}_0$. If the per-capita contact rate is not adjusted, a reduction in the size of the interacting population in an SIR model reduces per-capita contacts.

Modify SIR Eqs. (1)–(5) determining disease epidemiology within the subgroup as follows:

$$S_{tg} + I_{tg} + R_{tg} = w_g \tag{C.9}$$

$$\dot{I}_{tg} = \frac{\beta_g I_{tg} S_{tg}}{w_g} - (\alpha + \mu) I_{tg} \tag{C.10}$$

$$\dot{R}_{tg} = \alpha I_{tg} - \mu R_{tg} \tag{C.11}$$

$$\dot{S}_{tg} = w_g \mu - \frac{\beta_g I_{tg} S_{tg}}{w_g} - \mu S_{tg}. \tag{C.12}$$

This system of equations exhibits several differences from the original. First, the compartments evolve independently for each group so are indexed by the group g. Group size w_g and disease transmissibility β_g are also indexed by g. The remaining parameters (μ, α) are common across groups so not subscripted. Another difference is that there is no vaccine in this thought experiement, so $Q = 0$ and $Z_t = 0$. Another difference is that the population size on the right-hand side of (C.9) is w_g, not 1. This shows up again in (C.12) in that the inflow of newborn susceptibles is $w_g \mu$ not μ as before. The key difference is the division of the $\beta_g I_{tg} S_{tg}$ term by w_g in (C.10) and (C.12), in effect increasing the per-person contact rate in a smaller population so that total contacts remains constant for a person separated from a larger population into a subgroup.

Let $\mathscr{R}_{0g} = \beta_g/(\alpha + \mu)$. Setting $\dot{I}_{tg} = \dot{R}_{tg} = \dot{S}_{tg} = 0$ in (C.9)–(C.12) and solving for the steady-state infection rate yields

$$I_g = \max\left[0, \frac{\mu}{\alpha + \mu}\left(1 - \frac{1}{\mathscr{R}_{0g}}\right)\right] w_g, \tag{C.13}$$

implying

$$I_g = \frac{w_g \mu}{\alpha + \mu}\left(1 - \frac{1}{\mathscr{R}_{0g}}\right) \tag{C.14}$$

if $\mathscr{R}_{0g} > 1$.

Tables 2 and 3 state that $\text{MEX}^* = \theta H/\mathscr{R}_0$ under both perfect competition and monopoly for $\mathscr{R}_0 \in (1, 1/(1-\tilde{c})]$. In fact, one can show that the different values MEX^* takes on for yet higher $\mathscr{R}_0$ are due to changes in vaccine consumption behavior at higher $\mathscr{R}_0$, not the increase in $\mathscr{R}_0$ itself. Constraining $Q = 0$, which shuts down changes in consumer behavior, one can show via direct calculation that $\text{MEX}(0) = \theta H/\mathscr{R}_0$ for all $\mathscr{R}_0 > 1$. This relationship can also be shown to hold at the group level. Hence, for all $\mathscr{R}_{0g} > 1$,

$$\text{MEX}_g(0) = \frac{\theta H}{\mathscr{R}_{0g}}. \tag{C.15}$$

The researcher believes (C.13) applies to the population as a whole with a homogeneous basic reproductive number $\mathscr{R}_0$. Dropping the group subscript, setting $w_g = 1$, and assuming that the research observes a positive aggregate infection rate gives the following estimating equation for $\mathscr{R}_0$:

$$\mathscr{R}_0 = \left[1 - \left(\frac{\alpha+\mu}{\mu}\right)\sum_{g\in\Gamma} I_g\right]^{-1}. \tag{C.16}$$

Combining this with the relationship $\text{MEX}(0) = \theta H/\mathscr{R}_0$ that the researcher believes applies at the aggregate level yields the following estimating equation for the aggregate marginal externality:

$$\text{MEX}(0) = \theta H\left[1 - \left(\frac{\alpha+\mu}{\mu}\right)\sum_{g\in\Gamma} I_g\right] = \sum_{g\in\Gamma}\left(w_g\frac{\theta H}{\mathscr{R}_{0g}}\right) = \sum_{g\in\Gamma} w_g \text{MEX}_g(0). \tag{C.17}$$

The second equality follows from substituting from (C.14), and the last equality follows from substituting from (C.15). Thus, if $\mathscr{R}_{0g} > 1$ for all $g \in \Gamma$, the naïve estimate of aggregate $\text{MEX}(0)$ is the weighted arithmetic mean of $\text{MEX}_g(0)$.

Suppose $\mathscr{R}_{0g} \leq 1$ for $g \in \Gamma'$ and $\mathscr{R}_{0g} > 1$ for $g \in \Gamma''$, where $\Gamma', \Gamma'' \neq \emptyset$. Then $I_g = 0$ for $g \in \Gamma'$, implying (C.17) becomes

$$\text{MEX}(0) = \theta H\left[1 - \left(\frac{\alpha+\mu}{\mu}\right)\sum_{g\in\Gamma''} I_g\right] \tag{C.18}$$

$$= \theta H\left[1 - \sum_{g\in\Gamma''} w_g\left(1 - \frac{1}{\mathscr{R}_{0g}}\right)\right] \tag{C.19}$$

$$= \theta H\left(1 - \sum_{g\in\Gamma''} w_g\right) + \sum_{g\in\Gamma''} \text{MEX}_g(0) \tag{C.20}$$

$$= \theta H \sum_{g\in\Gamma'} w_g + \sum_{g\in\Gamma} \text{MEX}_g(0). \tag{C.21}$$

Eq. (C.21) shows that the researcher's estimate of $\text{MEX}(0)$ is biased upward by a term proportional to the share of the population in groups with $\mathscr{R}_{0g} \leq 1$. Implicit in the researcher's ability to compute $\text{MEX}(0)$ from an estimate of $\mathscr{R}_0$ is knowledge of θH. If the researcher also knows $\sum_{g\in\Gamma'} w_g$, an unbiased estimate of $\sum_{g\in\Gamma} \text{MEX}_g(0)$ can be computed simply by subtracting $\theta H \sum_{g\in\Gamma'} w_g$ from the naïve estimator $\text{MEX}(0)$.

In summary, the researcher obtains an unbiased estimate of the weighted arithmetic mean of MEX across groups but only if $\mathscr{R}_{0g} > 1$ for all $g \in \Gamma$. If $\mathscr{R}_{0g} < 1$ for some g, then the researcher will overestimate mean MEX. The discontinuity in the functional relationship between $\mathscr{R}_{0g}$ and MEX_g at $\mathscr{R}_{0g} = 1$ biases the calculation of the mean. The bias can be corrected by subtracting the product of θH and the proportion of consumers in subgroups g with $\mathscr{R}_{0g} \leq 1$ from the naïve estimate of MEX.

CRediT authorship contribution statement

Matthew Goodkin-Gold: Formal analysis. **Michael Kremer:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Supervision. **Christopher M. Snyder:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Heidi Williams:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Visualization.

References

Acemoglu, D., Chernozhukov, V., Werning, I., Whinston, M.D., 2021. Optimal targeted lockdowns in a multigroup SIR model. Am. Econ. Rev. Insights 3 (4), 487–502.

Acemoglu, D., Makhdoumi, A., Malekian, A., Ozdaglar, A., 2020. Testing, voluntary social distancing and the spread of an infection. National Bureau of Economic Research working paper no. 27483.

Adida, E., Dey, D., Mamani, H., 2013. Operational issues and network effects in vaccine markets. Eur. J. Oper. Res. 231 (2), 414–427.

Althouse, B.M., Bergstrom, T.C., Bergstrom, C.T., 2010. A public choice framework for controlling transmissible and evolving diseases. Proc. Natl. Acad. Sci. 107, 1696–1701. (supplement 1)

Alvarez, F., Argente, D., Lippi, F., 2021. A simple planning problem for COVID-19 lock-down, testing, and tracing. Am. Econ. Rev. Insights 3 (3), 367–382.

Anderson, R., May, R., 1991. Infectious Diseases of Humans. Oxford University Press.

Atkeson, A., Kopecky, K., Zha, T., 2020. Estimating and forecasting disease scenarios for COVID-19 with an SIR model. National Bureau of Economic Research working paper no. 27335.

Atkeson, A., Kopecky, K., Zha, T., 2021. Behavior and the transmission of COVID-19. Am. Econ. Assoc. Pap. Proc. 111, 356–360.

Avery, C., 2021. A simple model of social distancing and vaccination. National Bureau of Economic Research working paper no. 29463.

Avery, C., Bossert, W., Clark, A., Ellison, G., Ellison, S., 2020. An economist's guide to epidemiology models of infectious disease. J. Econ. Perspect. 34 (4), 79–104.

Bauch, C.T., Earn, D.J.D., 2004. Vaccination and the theory of games. Proc. Natl. Acad. Sci. 101 (36), 13391–13394.

Boulier, B., Datta, T., Goldfarb, R., 2007. Vaccination externalities. B. E. J. Econ. Anal. Policy 7 (1). Article 23

Brennan, M.J., 2009. The US food and drug administration provides a pathway for licensing vaccines for global diseases. PLoS Med. 6 (7), e1000095.

Brito, D.L., Sheshinski, E., Intrilligator, M.D., 1991. Externalities and compulsory vaccination. J. Public Econ. 45 (1), 69–90.

Chen, F., Toxvaerd, F., 2014. The economics of vaccination. J. Theor. Biol. 363, 105–117.

Chowell, G., Fenimore, P.W., Castillo-Garsow, M.A., Castillo-Chavez, C., 2003. SARS Outbreaks in ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism. J. Theor. Biol. 224, 1–8.

Diekmann, O., Heesterbeek, J.A.P., Metz, J.A.J., 1990. On the definition and the computation of the basic reproduction ratio r_0 in models for infections diseases in heterogeneous populations. J. Math. Biol. 28 (4), 365–382.

Eccleston-Turner, M., 2016. The economic theory of patent protection and pandemic influenza vaccines: do patents really incentivize innovation in the field? Am. J. Law Med. 42 (2–3), 572–597.

Eichenbaum, M. S., Rebelo, S., Trabandt, M., 2020. The macroeconomics of epidemics. National Bureau of Economics working paper no. 26882.

Ellison, G., 2020. Implications of heterogeneous SIR models for analyses of COVID-19. National Bureau of Economics working paper no. 27373.

Fajgelbaum, P.D., Khandelwal, A., Kim, W., Mantovani, C., Schaal, E., 2021. Optimal lockdown in a commuting network. Am. Econ. Rev. Insights 3 (4), 503–522.

Farboodi, M., Jarosch, G., Shimer, R., 2020. Internal and external effects of social distancing in a pandemic. National Bureau of Economic Research working paper no. 27059.

Fenichel, E.P., 2013. Economic considerations for social distancing and behavioral based policies during an epidemic. J. Health Econ. 32 (2), 440–451.

Finkelstein, A., 2004. Static and dynamic effects of health policy: evidence from the vaccine industry. Q. J. Econ. 119 (2), 527–564.

Francis, P.J., 1997. Dynamic epidemiology and the market for vaccinations. J. Public Econ. 63 (3), 383–406.

Funk, S., Salathé, M., Jansen, V.A.A., 2010. Modelling the influence of human behaviour on the spread of infectious diseases: a review. J. R. Soc. Interface 7, 1247–1256.

Galeotti, A., Rogers, B.W., 2013. Strategic immunization and group structure. Am. Econ. J. Microecon. 5 (2), 1–32.

Gans, J., 2020. The economic consequences of $\hat{R} = 1$: towards a workable behavioural epidemiological model of pandemics. National Bureau of Economic Research working paper no. 27632.

Geoffard, P.-Y., Philipson, T., 1997. Disease eradication: public vs. private vaccination. Am. Econ. Rev. 87 (1), 222–230.

Gersovitz, M., 2003. Births, recoveries, vaccinations, and externalities. chapter 2. In: Arnott, B., Greenwald, R.K., Nalebuff, B. (Eds.), Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz. MIT Press.

Gersovitz, M., Hammer, J.S., 2004. The economical control of infectious diseases. Econ. J. 114 (492), 1–27.

Gersovitz, M., Hammer, J.S., 2005. Tax/subsidy policy toward vector-borne infectious diseases. J. Public Econ. 89 (4), 647–674.

Goodkin-Gold, M., Kremer, M., Snyder, C. M., Williams, H., 2021. Optimal vaccine subsidies for epidemic diseases. University of Chicago working paper.

Jones, C. J., Philippon, T., Venkateswaran, V., 2020. Optimal mitigation policies in a pandemic: social distancing and working from home. National Bureau of Economic Research working paper no. 26984.

Kennedy, C.E., Yeh, P.T., Atkins, K., Fonner, V.A., Sweat, M.D., O'Reilly, K.R., Rutherford, G.W., Baggaley, R., Samuelson, J., 2020. Economic compensation interventions to increase uptake of voluntary medical male circumcision for HIV prevention: asystematic review and meta-analysis. PLoS ONE 15 (1), e0227623.

Keppo, J., Kudlyak, M., Quercioli, E., Smith, L., Wilson, A., 2021. The behavioral SI* model, with applications to the swine flu and COVID-19 pandemics. National University of Singapore working paper.

Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. Proc. R. Soc. A Math. Phys. Eng. Sci. 115 (772), 700–721.

Kremer, M., 1996. Integrating behavioral choice into epidemiological models of AIDS. Q. J. Econ. 111 (2), 549–573.

Kremer, M., Snyder, C.M., 2015. Preventives versus treatments. Q. J. Econ. 130 (3), 1167–1239.

Kremer, M., Snyder, C.M., 2018. Preventives versus treatments redux: tighter bounds on distortions in innovation incentives with an application to the global demand for HIV pharmaceuticals. Rev. Ind. Organ. 53 (1), 235–273.

Liu, Y., Gayle, A.A., Wilder-Smith, A., Rocklöv, J., 2020. The reproductive number of COVID-19 is higher compared to SARS coronavirus. J. Travel Med. 27 (7), 1–4.

Liu, Y., Rocklöv, J., 2021. The reproductive number of the delta variant of SARS-cov-2 is far higher compared to the ancestral SARS-cov-2 virus. J. Travel Med. 28 (7), 1–3.

Makris, M., Toxvaerd, F., 2020. Great expectations: social distancing in anticipation of pharmaceutical innovations. University of Cambridge Working Papers in Economics no. 2097.

Mamani, H., Adida, E., Dey, D., 2012. Vaccine market coordination using subsidy. IIE Trans. Healthc. Syst. Eng. 2 (1), 78–96.

Manfredi, P., D'Onofrio, A., 2013. Modeling the interplay between human behavior and the spread of infectious diseases. Springer-Verlag.

Manski, C.F., 2010. Vaccination with partial knowledge of external effectiveness. Proc. Natl. Acad. Sci. 107 (9), 3953–3960.

Manski, C.F., 2017. Mandating vaccination with unknown indirect effects. J. Public Econ. Theory 19 (3), 603–619.

Manski, C. F., 2021. Vaccination planning under uncertainty, with application to covid-19. National Bureau of Economic Research working paper no. 28446.

Martcheva, M., 2015. An introduction to mathematical epidemiology. Springer.

McAdams, D., 2020. Nash SIR: an economic-epidemiological model of strategic behavior during a viral epidemic. Social Science Research Network working paper no. 393272.

McAdams, D., 2021. The blossoming of economic epidemiology. Ann. Rev. Econ. 13, 539–570.

Peltzman, S., 1975. The effects of automobile safety regulation. J. Political Econ. 83 (4), 677–725.

Philipson, T., 2000. Economic Epidemiology and Infectious Disease. Chapter 33 in A. J. Culyer and J. P. Newhouse, eds., Handbook of Health Economics (vol. 1). Elsevier.

Rachel, L., 2020. An analytical model of covid-19 lockdown. London School of Economics Center for Macroeconomics discussion paper no. 2020–29.

Regula, T.C., 2010. Game theory of social distancing in response to an epidemic. PLoS Comput. Biol. 6 (5), e1000793.

Toxvaerd, F., 2019. Rational disinhibition and externalities in prevention. Int. Econ. Rev. 60 (4), 1737–1755. Philadelphia
Toxvaerd, F., 2020. Equilibrium social distancing. University of Cambridge Working Papers in Economics no. 2021.
Tröger, T., 2020. Optimal testing and social distancing of individuals with private health signals. University of Bonn discussion paper series CRC TR 224, no. 229.
U.S. Centers for Disease Control and Prevention, (2020). Vaccines and preventable diseases: Vaccines by disease. Hepatitis B vaccine efficacy downloaded from https://www.cdc.gov/vaccines/pubs/pinkbook/hepb.html, measles vaccine efficacy downloaded from https://www.cdc.gov/vaccines/vpd/mmr/public/index.html, pertussis and Tetanus vaccine efficacy downloaded from https://www.cdc.gov/vaccines/vpd/dtap-tdap-td/hcp/about-vaccine.htmlfor, polio vaccine efficacy downloaded from https://www.cdc.gov/vaccines/vpd/polio/index.html, and shingles vaccine efficacy downloaded from https://www.cdc.gov/vaccines/vpd/shingles/index.html.