Optimal Vaccine Subsidies for Epidemic Diseases

Matthew Goodkin-Gold
Harvard University Department of Economics

Michael Kremer
University of Chicago Department of Economics
and National Bureau of Economic Research (NBER)

Christopher M. Snyder
Dartmouth College Department of Economics and NBER

Heidi Williams
Stanford University Department of Economics and NBER

November 2021

Abstract: The positive epidemiological externality associated with vaccines provides a rationale for subsidies. We study how optimal subsidies vary with disease characteristics by integrating a standard epidemiological model into a vaccine market with rational economic agents. We focus on a vaccine campaign to quell an epidemic like Covid-19 in the short run. Across market structures ranging from competition to monopoly, we find that the infection rate, marginal externality, and optimal subsidy are nonmonotonic in transmissibility, peaking for diseases that spread quickly but not so quickly as to drive all consumers to become vaccinated. We study when universal vaccination emerges in equilibrium, when vaccination exhibits increasing social returns (providing an argument for concentrating a capacity-constrained campaign in few regions), and when producers' bias away from vaccines toward treatments generates deadweight loss. An appendix explores extensions to Cournot competition, consumer heterogeneity, and an alternative model of an endemic disease like measles in which new cohorts are continuously vaccinated in the steady state.

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

Keywords: vaccine, epidemiology, externality, pharmaceutical

Acknowledgements: The authors are grateful for helpful comments from Witold Więcek and seminar participants in the Harvard Economics Department, Yale School of Medicine, the "Infectious Diseases in Poor Countries and the Social Sciences" conference at Cornell University, the DIMACS "Game Theoretic Approaches to Epidemiology and Ecology" workshop at Rutgers University, the "Economics of the Pharmaceutical Industry" roundtable at the Federal Trade Commission's Bureau of Economics, the U.S. National Institutes of Health "Models of Infectious Disease Agent" study group at the Hutchinson Cancer Research Center in Seattle, the American Economic Association "Economics of Infectious Disease" session, the Health and Pandemics (HELP!) Economics Working Group "Covid-19 and Vaccines" workshop, and the NBER "COVID-19 and Health Outcomes" conference. Christopher Cardillo, Maya Durvasula, Nishi Jain, Amrita Misha, Frank Schilbach, and Alfian Tjandra provided excellent research assistance. Williams gratefully acknowledges financial support from NIA grant number T32-AG000186 to the NBER.

1. Introduction

Technologies that help prevent infectious diseases such as vaccines, social distancing, condoms, and mosquito nets can generate positive health externalities. While standard economic models provide a justification for public subsidies of such preventive technologies, existing models provide little guidance on the appropriate magnitude of these subsidies. This gap in understanding makes it difficult for economists to provide guidance—even at a conceptual level—on both the optimal level of government subsidies for infectious disease control and the ways in which the level of such subsidies should vary across diseases. Furthermore, most existing work by economists has been oriented toward endemic diseases and policies which play out over years or decades such as vaccination campaigns to eradicate polio or circumcision to reduce the spread of HIV. However, the Covid-19 pandemic has underscored the urgency of understanding short-run policy responses to the epidemics that can result from the emergence of novel diseases or from localized outbreaks of otherwise dormant diseases such as Ebola or MERS.

To address these questions, we construct a tractable model integrating epidemiological and economic considerations. For concreteness, we focus on the market for a vaccine, but the analysis applies to other aforementioned preventive technologies. Consumers and producers base their economic decisions on rational expectations of disease dynamics based on a susceptible-infected-recovered (SIR) model standard in the epidemiology literature. Most of the paper adopts a short-run perspective of a vaccine campaign introduced at a single point in time into an SIR epidemic without population turnover. A complemenary long-run perspective, suited to an endemic disease, is provided by an extension in which we incorporate population turnover into the SIR model and analyze steady-state vaccination rates.

A key finding—which holds in both the short- and long-run analysis, across market structures ranging from perfect competition to Cournot to monopoly, and for homogeneous or heterogeneous consumers—is that the marginal externality of one's vaccination on others is nonmonotonic in the disease's basic reproductive ratio $\mathcal{R}_0$, a widely used measure of infectiveness. For low values of $\mathcal{R}_0$, the marginal externality is low because there is little disease transmission between people. For high values of $\mathcal{R}_0$, vaccinating a given consumer does not provide much protection to others since they are almost certain to contract the disease from another source anyway. To be sure, a consumer's vaccination provides a substantial social benefit when $\mathcal{R}_0$ is extremely high, but most of that benefit is internalized by the consumer. The marginal externality thus peaks for intermediate values of $\mathcal{R}_0$.

Surprisingly, the uninternalized externality can be so large that the infection rate is hump-shaped in $\mathcal{R}_0$. The optimal vaccine subsidy is always hump-shaped in $\mathcal{R}_0$.

Going beyond nonmonotonicies, Section 3.2 explores conditions under which universal vaccination can be a viable business strategy. Previous game-theoretic analyses of vaccine uptake pointed out that a perfectly effective vaccine would never be universally purchased at a positive price because, with all other consumers protected, the marginal consumer obtains no private benefit (Geoffard and Phillipson 1997, May 2000, Bauch and Earn 2004). In our short-run analysis, however, universal vaccination with a perfectly effective vaccine can be profitable. The risk of contracting the disease from those infected before the arrival of the vaccine but not yet recovered preserves a positive willingness to pay for the marginal consumer even if all other susceptibles are protected. Universal vaccination with a perfectly effective vaccine is not just a possible equilibrium for some parameters for some market structure, it is guaranteed in equilibrium even under monopoly for sufficiently low cost and sufficiently high infectiousness.

Section 5 shows that if the product of $\mathcal{R}_0$ and $\hat{S}_0$ (the susceptible proportion of the population upon vaccine rollout) exceeds 2, social returns to vaccination are initially increasing, implying that a small capacity would be more efficiently concentrated than spread evenly across regions. The condition $\mathcal{R}_0\hat{S}_0 > 2$ has intuitive appeal. Infectiousness $\mathcal{R}_0$ promotes disease spread much as prime weather conditions (low moisture, hot temperature, high winds) promote a forest fires' spread; the stock of susceptibles $\hat{S}_0$ is like flammable material providing fuel for the fire. For an epidemic to grow initially rather than diminishing from the outset requires a certain level of both factors in the absence of a vaccine, technically $\mathcal{R}_0\hat{S}_0 > 1$. When $\mathcal{R}_0\hat{S}_0$ exceeds the higher threshold of 2, the epidemic is so explosive that a small amount of vaccine does little to slow it; to make a measurable dent in the epidemic requires concentrating supplies in one region. If $\mathcal{R}_0\hat{S}_0$ exceeds a yet higher threshold than 2 that we specify, vaccination exhibits increasing social returns for all capacity levels, meaning that would be efficient to serve all susceptibles in a region before moving to the next region.

Section 6 compares the results to a market for a drug is similar in all ways to the vaccine except that it treats symptoms but does nothing to reduce disease spread from treated individuals. We show that a monopolist would always prefer to develop the drug but parameters exist for which social welfare is higher with the vaccine. Consistent with nonmonotonicities found elsewhere, the monopolist's bias toward a drug is greatest for intermediate values of $\mathcal{R}_0$.

For illustrative purposes, we provide stylized calibrations of both short- and long-run analyses to real-world diseases. The featured calibration for the short-run analysis is to the Covid pandemic.

We find that the vaccine market achieves the first best under perfect competition but not under monopoly, resulting in a deadweight loss of more than quarter of first-best welfare. To deliver the first best, a subsidy of more than half of the monopoly price is required. Social returns to a Covid vaccine are initially increasing but only mildly so. The featured calibrations for the long-run analysis are to HIV and measles. Optimal subsidies may be very large relative to current levels in the case of competitively supplied products. For example, the external benefit from using circumcision to prevent HIV in developing countries could justify subsidies of more than $1,000, while programs studied to date paid participants no more than $15 beyond procedure costs. For monopoly products sold directly to consumers, a per-dose subsidy may not always be a viable policy option since the minimum subsidy achieving the first best can be enormous: in our measles calibration, for example, fifteen times the harm from actually having the disease. Given plausible values for the social cost of public funds, such subsidies would be prohibitively expensive. Bulk purchase of the optimal quantity coupled with subsidized distribution may be a cheaper route to the first best.

While the SIR model we use is standard in the epidemiology literature by design, our novel contribution to that literature is to incorporate an epidemiological model into a welfare economics framework, facilitating the analysis of market equilibrium and optimal policy. Much of the epidemiology literature focuses on characterizing when disease eradication is feasible, whereas our framework admits nontrivial analysis even in settings in which eradication is impossible (long-run equilibrium with a perfectly effective vaccine) or certain (the disease always dies out in our short-run analysis). An important exception is Althouse, Bergstrom, and Bergstrom (2010), who also consider welfare analysis of vaccination, calibrating a simple model for four prominent diseases to estimate optimal subsidies under perfect competition and perfectly effective vaccination. Our paper builds on their work, allowing for imperfect vaccines, including a supply-side model of firm behavior, and generating comparative statics which allow theoretical insights into how epidemiological and economic parameters impact market outcomes and optimal policy.

The epidemiology literature previously recognized the possibility that the nonlinear nature of epidemics may dictate optimal policy concentrating a scarce stockpile in one population rather spreading across them. Keeling and Shattock (2012) provided an early contribution, subsequently refined by work including Keeling and Ross (2015), Nguyen and Carlson (2016), and Enayati and Özaltin (2020). This literature has the advantage of studying increasingly rich epidemiological models, the results are simulated in numerical examples. We contribute a formal conceptualization of initial and eventual increasing social returns and aid understanding by providing a necessary and

sufficient condition for these outcomes in analytical form.

During the current Covid pandemic, scholars have sought to apply detailed models to forecast the course of the pandemic (e.g., Atkeson, Kopecky, and Zha 2020) and to recommend policies for prioritizing scarce vaccine supplies among heterogeneous consumers (e.g., Buckner, Chowell, and Springborn 2021). The context of our Covid calibration is different—on optimal subsidies in a decentralized market rather than optimal strategies for a central planner—as is our goal—obtaining qualitative results in a stylized model rather than quantitative results in a more complex model. Recognizing, among other market failures, the free-rider problem we study here, economists have advocated direct government funding of expanded capacity for Covid vaccines (Ahuja *et al*. 2021, Castillo *et al*. 2021), which most countries are currently providing at no charge to their citizens. Given the potential role of social distancing in quelling the pandemic, a contemporaneous literature has sought to integrate endogenous social distancing into epidemiological models (Eichenbaum, Rebelo, and Trabandt 2020; Farboodi, Jarosch, and Shimer 2020; Gans 2020; Jones, Philippon, and Venkateswaran 2020; McAdams 2020; Rachel 2020; Toxvaerd 2020).[1] This literature, contemporaneous with our work, studies what in effect is a competitively-supplied technology, whereas we examine various market structures. We have a different focus, characterizing the nonmonotonicity of externalities and optimal subsidies as a function of disease infectiveness.

Economists have long observed that vaccines may provide positive externalities that could affect consumers' and firms' decisions (see, among others, Brito, Sheshinski, and Intrilligator 1991; Chen and Toxvaerd 2014; Francis 1997; Geoffard and Philipson 1997; Gersovitz 2003; Gersovitz and Hammer 2004, 2005).[2] Boulier, Datta, and Goldfarb (2007) use a standard epidemiological model alone (i.e., neither interacted with consumer decisions nor a supply-side model of firm behavior) to examine properties of vaccination externalities that arise solely due to epidemiological concerns. Geoffard and Philipson (1997) use an epidemiological model similar to ours to show that a vaccine producer with market power will not choose to eradicate the disease in the steady-state. Galeotti and Rogers (2013) model vaccination choices in a heterogenous population, and consider the effect of network structures in determining optimal vaccine allocation.[3] Economists have attempted to

[1]Earlier work on social distancing and other behavioral responses to epidemics includes Kremer (1996), Reluga (2010), Fenichel (2013), and Toxvaerd (2019).

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

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

provide well-identified empirical estimates of vaccine externalities (Cook *et al.* 2009, Ward 2014, Bethune and Korinek 2020) among yet more ambitious attempts to structurally estimate an epidemic model with behavioral responses (Greenwood *et al*. 2019, Aguirregabiria *et al.* 2020, Bisin and Moro 2020). Given the difficulty of using randomized controlled trials to estimate externalities, Manski's (2010, 2017, 2021) articles provide theoretical guidance on optimal vaccine policies (including mandates) when the extent of externality is unknown. We make a number of contributions to this economics literature. First, we incorporate a model of firm behavior and explicitly characterize equilibrium solutions for both positive and normative outcomes—specifically, externalities and optimal subsidies—in terms of estimable parameters. Second, our short-run analysis provides what to our knowledge is the first treatment of vaccination externalities and subsidies oriented to the initial stages of a novel disease or localized outbreak and the first to characterize increasing returns to vaccination in this setting.

Our paper is perhaps closest to Mamani, Adida, and Dey (2012) and Adida, Dey, and Mamani (2013) in the operations-research literature. Their insightful papers also analyze optimal subsidies, nesting endemic and epidemic cases, allowing for a general Cournot market structure. Their focus is on consumers with uniformly distributed harm. While we also examine consumer heterogeneity (see Online Appendix B3), our analysis focuses on homogeneous consumers, offering several advantages. Their results are left in terms of reduced-form functions amalgamating equilibrium vaccine coverage and $\mathcal{R}_0$. Our more definitive expressions afford additional insights, most importantly allowing us to analyze the comparative-static effect of increases in $\mathcal{R}_0$. Our central result on the nonomontonicity of optimal subsidies in $\mathcal{R}_0$ and other comparative statics are new in our paper. Another advantage of the homogenous-consumer model is to clarify the relationship between consumers' marginal private benefits and vaccine coverage. An increase in coverage reduces marginal private benefit solely through the externality of interest and is not an artifact a move down the exogenously imposed distribution of consumer harm. Solving for equilibrium in the homogeneous-consumer does present a more challenging fixed-point problem, which is a methodological contribution of our paper. We also provide calibrations, results on increasing social returns, and a comparison between drugs and vaccines not found in their papers.

2. Model

We begin with the short-run analysis of a market for a vaccine against a disease that rises and falls as an epidemic in a shorter spell than a human generation, modeling a campaign that rolls out the vaccine to substantial fraction of the population quickly to mitigate the harm experienced by the current generation, with future generations spared much damage by the epidemic's natural decline. We begin with this short-run analysis and feature it throughout most of the paper because, first, it is relevant for the Covid pandemic of current interest and, second, requires more discussion given the more delicate mathematics involved. For pedagogical purposes, in this variant of the model, we adopt the extreme assumption that all doses of vaccine that will ever be administered are administered in a single instant.

2.1. Epidemiology

The foundation of our analysis is the standard susceptible-infected-removed (SIR) epidemiological model due to Kermack and McKendrick (1927). Continuous time indexed by t begins with the arrival of the vaccine at date $t = 0$. Assume for simplicity that the disease is non-fatal and that there are no births or deaths within the short time frame considered, leaving the population size constant over time. Anticipating their role in the vaccine market, we call members of this population "consumers." Throughout most of the analysis, we assume consumers are homogeneous in harm, disease spread, and all other dimensions.

Consumers are partitioned into four compartments: susceptible to infection S_t, currently infected I_t, recovered from an infection R_t, or immunized V_t. Normalizing the population mass to 1,

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

the compartments can then be interpreted either as masses or proportions. Compartments evolve according to the following equations:

$$\dot{S}_t = -\beta I_t S_t \tag{2}$$

$$\dot{I}_t = \beta I_t S_t - \alpha I_t \tag{3}$$

$$\dot{R}_t = \alpha I_t \tag{4}$$

$$\dot{V}_t = 0. \tag{5}$$

Dots over variables denote derivatives with respect to t; e.g., $\partial S_t/\partial t \equiv \dot{S}_t$.

A susceptible consumer is assumed to contract 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. Equation (2) indicates that the susceptible population falls by the number of newly infected. Equation (3) indicates that the infected population is increased by the number of newly infected and reduced by the mass αI_t of previously infected consumers who recover, where $\alpha \in (0,1)$ denotes the recovery rate. This αI_t mass flows into R_t, as indicated by equation (4). Under the assumption that recovered individuals cannot be reinfected, this is the only change to R_t. Equation (5) reflects the instantaneous nature of the vaccination campaign, with no further vaccine administered after the initial tranche at date 0.[4] We assume that if the initial dose is not effective for a person, further doses will not be effective for that person either. Under that assumption, administering all vaccine in the first instant is both the profit-maximizing and welfare-maximizing strategy.[5]

Let Q denote the vaccine quantity administered at date 0. 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. Let $\theta \in (0,1)$ denote vaccine efficacy. Let $\hat{S}_0$, $\hat{I}_0$, and $\hat{R}_0$ denote the counterfactual value of the relevant compartments at date 0 in the absence of a vaccine (by definition, $\hat{V}_0 = 0$). Then the initial conditions for the SIR system can be written

$$S_0 = \hat{S}_0 - \theta Q \tag{6}$$

$$I_0 = \hat{I}_0 \tag{7}$$

$$R_0 = \hat{R}_0 = 1 - \hat{I}_0 - \hat{S}_0 \tag{8}$$

$$V_0 = \theta Q. \tag{9}$$

Note that V_0 is the product of vaccinations Q and efficacy θ since V_t includes only successfully immunized consumers. Given that susceptibles are the only individuals that can possibly benefit

[4]Epidemiology texts label the vaccination process involved in our short-run analysis "vaccination at recruitment" (Martcheva 2015, Section 9.2.1), contrasting the process of "continuous vaccination" involved in our long-run analysis (Martcheva 2015, Section 9.2.2).

[5]Logistical constraints would prevent such rapid vaccine rollout in practice, but the model may still be a reasonable approximation to an intensive vaccine campaign against Covid or other epidemic disease.

from vaccination, in equilibrium,

$$Q \leq \hat{S}_0, \tag{10}$$

ensuring $S_0 \geq 0$. We treat $\hat{I}_0$ and $\hat{S}_0$ as exogenous parameters, allowing them to take on any admissible values $\hat{I}_0 \in (0,1)$, $\hat{S}_0 \in (0, 1-\hat{I}_0]$.

In lieu of the transmission parameter β, epidemiologists often work with a related parameter called the basic reproductive ratio $\mathcal{R}_0$, traditionally defined as the expected number of secondary infections generated by adding an infected individual to a fully susceptible population.[6] One can see that the disease eventually dies out in an unvaccinated population if $\mathcal{R}_0 < 1$ and remains endemic if $\mathcal{R}_0 > 1$. In our model,

$$\mathcal{R}_0 = \frac{\beta}{\alpha}. \tag{11}$$

To understand this expression, each instant the individual remains infected, he or she infects a number of others equal to β times the size of the susceptible population, which is approximately 1 during the period since the infected individual is introduced into a fully susceptible population. The individual remains infected for an expected duration of $1/\alpha$.[7] The subsequent analysis takes $\mathcal{R}_0$ as the key exogenous parameter, capturing the disease's infectiveness. Estimates of $\mathcal{R}_0$ vary considerably across diseases—from 1.1 for SARS (Chowell *et al.* 2003) at the low end to 16–18 for measles and pertussis at the high end (Anderson and May 1991)—as well as across time and region.

2.2. Preliminary Epidemiological Results

In subsequent notation, Q is appended as an argument to equilibrium variables to emphasize their dependence on that key variable to be endogenized later. Limiting compartment values at the end of the epidemic are denoted by $S_\infty(Q)$, $I_\infty(Q)$, and $R_\infty(Q)$; for example, $S_\infty(Q) = \lim_{t\uparrow\infty} S_t(Q)$.

The following series of lemmas characterize $S_t(Q)$ and $I_t(Q)$ both for finite and limiting values of t. The other compartments then follow easily from preceding equations: $V_t(Q) = \theta Q$ by (5) and (9) and $R_t(Q) = 1 - I_t(Q) - S_t(Q) - V_t(Q)$ by (1). We provide these preliminary results up front in the model section since they will be useful in fleshing out economic aspects of the model below. Appendix A provides proofs.

[6]The modern definition of $\mathcal{R}_0$ due to Diekmann, Heersterbeek, and Metz (1990) is the dominant eigenvalue of the next-generation operator in the epidemiological system. Martcheva (2015, p. 51) shows that equation (B33) provides the value of $\mathcal{R}_0$ implied by this definition.

[7]To see this, note that the sole risk of exiting the infected state is recovery, with hazard $\lambda_R(t) = \alpha$. In a Poisson duration models, the duration of a spell equals the reciprocal of the hazard, here $1/\lambda_R(t) = 1/\alpha$.

Lemma 1. *$I_t(Q) > 0$ and $S_t(Q) > 0$.*

Lemma 2. *$S_t(Q)$ is strictly decreasing in t.*

Lemma 3. *If $\mathcal{R}_0 S_0(Q) \leq 1$, then $I_t(Q)$ is strictly decreasing in t for all $t > 0$. Otherwise, $I_t(Q)$ is hump-shaped, peaking at time $T > 0$ satisfying $S_T(Q) = 1/\mathcal{R}_0$, strictly increasing for $t < T$, and strictly decreasing for $t > T$.*

Lemma 4. *The limits $I_\infty(Q)$ and $S_\infty(Q)$ exist. In particular, $I_\infty(Q) = 0$ and $S_\infty(Q) \in (0, S_0(Q))$.*

Lemma 5. *$S_\infty(Q) < 1/\mathcal{R}_0$.*

To explain the lemmas in intuitive terms, they tell us that the infection rate is always positive in finite time because, if not increasing, infections are at worst declining at a proportional rate less than 100% each instant, which can never force the infection rate to 0. The infection rate does asymptote to 0 as the stock of susceptibles is depleted and recovery takes over as the dominating force, reducing the stock of infecteds. Turning to results for the population of susceptibles, with an imperfectly effective vaccine ($\theta < 1$), even a universal vaccination campaign cannot eliminate the stock of susceptibles at date 0. The stock of susceptibles is never forced to 0 after because the proportional decline is less than 100% each instant. The stock of susceptibles strictly decreases over time since it is subject to outflows but not inflows.

According to Lemma 3, there are two possible shapes for infection rate over the ex post period. One possibility is that the infection rate continuously diminishes from its initial level. Another possibility is that infections have a hump-shaped path, expanding up to a peak and declining thereafter. For infections to initially expand rather than contract requires both infectiveness (i.e., high $\mathcal{R}_0$) and substantial fuel (i.e., high $S_0(Q)$). We will call $\mathcal{R}_0 S_0(Q) > 1$ the condition for epidemic expansion.

While no closed-form solution exists for the limiting number of susceptibles $S_\infty(Q)$, the next lemma, proved in Appendix A, expresses it as an implicit function of other model parameters (though the notation continues to emphasize only the variable to be endogenized later, Q, as an argument of S_∞). The lemma also provides an expression for $S_\infty(Q)$ in terms of the principal branch of the Lambert W function, here denoted $\bar{L}$.[8]

[8]The Lambert W function L frequently arises in epidemiological applications. By definition $L(x)$ is the implicit solution to the exponential equation $L(x)e^{L(x)} = x$. The principal branch $\bar{L}$ is the sole solution to the implicit equation or, if two solutions exist, the higher of the two. The lower branch $\underline{L}$ is defined when two solutions exist as the lower of the two. Though $\bar{L}$ and $\underline{L}$ do not have a closed-form solutions, they can be computed with built-in functions included in Matlab, R, and other standard software packages.

Lemma 6. *$S_\infty(Q)$ satisfies*

$$\ln S_\infty(Q) - \mathcal{R}_0 S_\infty(Q) = \ln(\hat{S}_0 - \theta Q) - \mathcal{R}_0(\hat{I}_0 + \hat{S}_0 - \theta Q) \tag{12}$$

and can be written

$$S_\infty(Q) = \frac{1}{\mathcal{R}_0}\left|\bar{L}\left(-\mathcal{R}_0(\hat{S}_0 - \theta Q)e^{-\mathcal{R}_0(\hat{I}_0+\hat{S}_0-\theta Q)}\right)\right|. \tag{13}$$

Equations (12) and (13) can be used to derive the limiting number steady-state population of susceptibles in the extremes of uninfective or infinitely infective diseases. If a disease cannot be transmitted, initial susceptibles never contract the disease, so remain as susceptibles in the steady state, implying $\lim_{\mathcal{R}_0\downarrow 0} S_\infty(Q) = S_0(Q) = \hat{S}_Q - \theta Q$, as can be proved by substituting $\mathcal{R}_0 = 0$ into (12).

Lemma 7. $\lim_{\mathcal{R}_0\downarrow 0} S_\infty(Q) = \hat{S}_0 - \theta Q$ *and* $\lim_{\mathcal{R}_0\downarrow 0}[\mathcal{R}_0 S_\infty(Q)] = 0$.

The second limit, recorded for reference, follows from $\lim_{\mathcal{R}_0\downarrow 0}[\mathcal{R}_0 S_\infty(Q)] = (\hat{S}_0 - \theta Q)\lim_{\mathcal{R}_0\downarrow 0}\mathcal{R}_0 = 0$.

In the opposite extreme of an infinitely infective disease, all susceptibles eventually become infected, implying that the steady-state susceptible population vanishes in the limit. This result can be derived formally from Lemma 5: $0 \le \lim_{\mathcal{R}_0\uparrow\infty} S_\infty(Q) \le \lim_{\mathcal{R}_0\uparrow\infty}(1/\mathcal{R}_0) = 0$, implying $\lim_{\mathcal{R}_0\uparrow\infty} S_\infty(Q) = 0$. A more subtle question regards the rate at which the steady-state susceptible population vanishes. The next lemma, proved in Appendix A, states that $S(Q)$ vanishes faster than $\mathcal{R}_0$ increases.

Lemma 8. $\lim_{\mathcal{R}_0\uparrow\infty} S_\infty(Q) = \lim_{\mathcal{R}_0\uparrow\infty}[\mathcal{R}_0 S_\infty(Q)] = 0$.

Comparative-static results can be obtained by applying the Implicit Function Theorem to (12). For instance,

$$\frac{\partial S_\infty(Q)}{\partial Q} = \frac{\theta S_\infty(Q)}{S_0(Q)}\left[\frac{\mathcal{R}_0 S_0(Q) - 1}{1 - \mathcal{R}_0 S_\infty(Q)}\right]. \tag{14}$$

Since its denominator is positive by Lemma 5, the sign of (14) is determined by whether the condition for epidemic expansion holds. The immediate effect of an increase in Q is to move an individual from the currently susceptible to the vaccinated compartment. If infections are waning, this immediate effect persists as a reduction in susceptibles that remain in the steady state. If the disease meets the condition for an epidemic, however, the reduction in current susceptibles has such a strong feedback effect in the form of reduced "fuel" for infections that a greater proportion of susceptibles remain in the steady state despite the immediate reduction in susceptibles.

2.3. Consumer Demand

Consumers are homogeneous and risk neutral. Consistent with the present short-run perspective, assume agents do not discount the future. Let H denote the total expected harm suffered by a consumer who contracts the disease over the spell before recovery.

The $\hat{S}_0$ individuals in the susceptible compartment when the vaccine is introduced are potential consumers. They make their demand decisions by comparing the vaccine's price P to their marginal private benefit, which can be written $MPB(Q) = \theta H \Phi_I(Q)$, where $\Phi_I(Q)$ denotes the probability a susceptible contracts the disease during the epidemic.

To compute $\Phi_I(Q)$, note that the probability an unvaccinated individual does not contract the disease equals $S_\infty(Q)/S_0(Q)$, the number of people who remain susceptible over the model's horizon divided by the number of people who are susceptible at the start of the ex post period. The probability of infection is the complementary probability

$$\Phi_I(Q) = 1 - \frac{S_\infty(Q)}{S_0(Q)} = 1 - \frac{S_\infty(Q)}{\hat{S}_0 - \theta Q}, \tag{15}$$

which Lemma 4 guarantees is positive. Thus,

$$MPB(Q) = \theta H \left[1 - \frac{S_\infty(Q)}{\hat{S}_0 - \theta Q} \right]. \tag{16}$$

Differentiating and rearranging yields

$$\frac{\partial MPB(Q)}{\partial Q} = \frac{-\theta \mathcal{R}_0 S_\infty(Q) MPB(Q)}{S_0(Q)[1 - \mathcal{R}_0 S_\infty(Q)]}, \tag{17}$$

which is negative by Lemma 5, confirming the intuition that vaccinating more consumers lowers their marginal private benefit.

Proceeding to derive the demand curve, all $\hat{S}_0$ consumers purchase the vaccine if $P < MPB(\hat{S}_0)$, and none purchase if $P > MPB(0)$. For P strictly between $MPB(\hat{S}_0)$ and $MPB(0)$, some but not all consumers purchase. Given they are homogeneous, consumers must be indifferent between purchasing and not, implying $P = MPB(Q)$. Given they are indifferent, any fraction of them are willing to purchase in equilibrium; demand is pinned down by the value of Q satisfying (16) when the right-hand side is set equal to P. Rearranging the resulting equation yields $S_\infty(Q) = (1 - P/\theta H)(\hat{S}_0 - \theta Q)$. Substituting this into (12) and solving for Q gives the following expression for demand when a

subset purchase:

$$d(P) = \frac{1}{\theta}\left\{\hat{S}_0 + \frac{\theta H}{P}\left[\frac{1}{\mathcal{R}_0}\ln\left(1-\frac{P}{\theta H}\right)+\hat{I}_0\right]\right\}. \tag{18}$$

Combining these facts yields the demand curve

$$D(P) = \begin{cases} 0 & P > MPB(0) \\ d(P) & P \in [MPB(\hat{S}_0), MPB(0)] \\ \hat{S}_0 & P < MPB(\hat{S}_0). \end{cases} \tag{19}$$

Equivalently, the demand curve is given by $d(P)$ unless this violates the boundary condition $d(P) \in [0, \hat{S}_0]$, in which case demand is given by the violated boundary.

2.4. Firm Supply

We analyze two different market structures in the text: perfect competition and monopoly. Appendix A provides results from a more general model of Cournot competition among n firms that nests these extremes.

Assume firms produce at constant marginal and average cost $c > 0$ per vaccine course (where a course involves multiple doses when needed to provide immunity). Under perfect competition, vaccine supply is perfectly elastic at price c. Under monopoly, the firm sets a price maximizing industry profit Π from sales made at date 0.

By equation (15) and Lemma 2, $\Phi_I(Q) < 1$, implying $MPB(Q) < \theta H$ by (16). There are no sales under perfect competition or indeed under any market structure if $c \geq \theta H$. To rule out trivial cases, throughout the remainder of the paper we assume

$$\frac{c}{\theta H} = \tilde{c} < 1, \tag{20}$$

introducing $\tilde{c}$ as shorthand notation to streamline subsequent expressions.

2.5. Normative Measures

Total harm experienced by consumers from the disease equals $HR_\infty(Q)$. Social benefit $SB(Q)$ is the complement of this, the harm avoided in the population who never contract the disease:

$$SB(Q) = H[1 - R_\infty(Q)] = H[S_\infty(Q) + \theta Q], \tag{21}$$

where the second equality follows from equation (1) and Lemma 4. Welfare $W(Q)$ is the difference between total social benefit and total vaccine production costs:

$$W(Q) = SB(Q) - cQ. \tag{22}$$

Marginal social benefit is the derivative $MSB(Q) = \partial SB(Q)/\partial Q$. Differentiating (21), substituting from (14) and (15), and rearranging yields

$$MSB(Q) = \frac{\theta H \Phi_I(Q)}{1 - \mathcal{R}_0 S_\infty(Q)}. \tag{23}$$

Let $MEX(Q) = MSB(Q) - MPB(Q)$ denote the marginal externality from a vaccine course. Substituting from (16) and (23) into this equality yields

$$MEX(Q) = \frac{\theta H \Phi_I(Q) \mathcal{R}_0 S_\infty(Q)}{1 - \mathcal{R}_0 S_\infty(Q)}. \tag{24}$$

Combining (15), (16), and (24) yields an equivalent expression for the marginal externality providing some useful intuition:

$$MEX(Q) = \mathcal{R}_0 S_\infty(Q) MSB(Q) = \mathcal{R}_0 S_0(Q) \left[\frac{S_\infty(Q)}{S_0(Q)}\right] MSB(Q). \tag{25}$$

The external benefit from vaccinating a given individual equals the social benefit of vaccinating everyone who would not be infected but for their interaction with that individual. By definition of the basic reproductive ratio, the given individual causes $\mathcal{R}_0 S_0(Q)$ direct infections in the susceptible population. However, some of those would have been infected by someone else later; only the fraction $S_\infty(Q)/S_0(Q)$ would have survived as susceptibles to the end but for their interaction with the given individual. The marginal externality equals the marginal social benefit cumulated over these "but for" infections.

Let Q^{**} denote the first-best quantity, maximizing $W(Q)$. If Q^{**} is not a corner solution, involving universal vaccination, it is an interior solution solving the social planner's first-order condition $MSB(Q^{**}) = c$.

3. Equilibrium

This section provides the analytical results for the model of a short-run epidemic. We present the solution for equilibrium under perfect competition and monopoly and characterize optimal government subsidies under those market structures. We then explore increasing social returns from vaccination and compare vaccines versus similarly effective treatments. The section concludes by providing results for relevant limiting parameter values.

3.1. Perfect Competition

Equilibrium values of variables will be distinguished with stars with an added subscript indicating the relevant market structure. Under perfect competition, the equilibrium price is $P_c^* = c$ and profit is $\Pi_c^* = 0$. The remaining equilibrium variables can be computed using straightforward algebra applied to the supplied equations. Table 1 reports the equilibrium values of selected variables as a function of $\mathcal{R}_0$. Appendix A provides derivations of the table entries.

The table distinguishes three relevant cases corresponding to three intervals for $\mathcal{R}_0$. In case (SR1), $\mathcal{R}_0$ is so low that no consumer finds it worthwhile to purchase the vaccine. The moderate values of $\mathcal{R}_0$ in case (SR2) lead some but not all susceptibles to purchase. The threshold dividing cases (SR1) and (SR2) is the value of $\mathcal{R}_0$ for which the interior solution for demand in (18) equals 0 (since no consumers purchase) at the equilibrium price $P_c^* = c$. Solving $d(c) = 0$ for $\mathcal{R}_0$ yields $\mathcal{R}_0 = |\ln(1-\tilde{c})|/(\hat{I}_0 + \tilde{c}\hat{S}_0)$. In the remaining cases, $\mathcal{R}_0$ is so high that all susceptibles find purchasing the vaccine worthwhile. The first best is obtained in these cases: $Q_c^* = Q^{**}$. The threshold dividing cases (SR2) and (SR3) is the value of $\mathcal{R}_0$ for which (18) equals $\hat{S}_0$ (since all susceptibles purchase) at $P_c^* = c$. Solving $d(c) = \hat{S}_0$ for $\mathcal{R}_0$ yields $\mathcal{R}_0 = |\ln(1-\tilde{c})|/[\hat{I}_0 + (1-\theta)\tilde{c}\hat{S}_0]$.

To visualize how the variables in Table 1 vary with $\mathcal{R}_0$, Figure 1 graphs a selection of them, one per panel, as functions of $\mathcal{R}_0$. Although some equilibrium variables do not have closed-form expressions, numerical methods can be used to generate the graphs. Focus for now on the dotted curves representing equilibrium under perfect competition. Vaccine quantity Q_c^*, graphed in the first panel, rises throughout case (SR2) from its value of 0 in case (SR1) to the first-best value Q^{**} in cases (SR3) and (SR4). It is unsurprising that equilibrium quantity is weakly increasing in the infectiveness of the disease measured by $\mathcal{R}_0$. Other equilibrium variables also display expected comparative statics in $\mathcal{R}_0$. MPB_c^* is weakly increasing and W_c^* is weakly decreasing in $\mathcal{R}_0$. It is noteworthy that MPB_c^* levels off at c in case (SR2). Given that some but not all consumers

purchase in this case, consumers must be indifferent between purchasing and not, implying that the equilibrium price $P_c^* = c$ must extract the entire marginal private benefit, implying $MPB_c^* = c$ over the whole interval.

Other variables display interesting nonmonotonicities. Starting from 0, the mass of people ever infected during the epidemic $R_\infty(Q_c^*)$ increases throughout case (SR1) due to the epidemiological effects of the higher $\mathcal{R}_0$ with no vaccine purchases in that case to offset it. In case (SR2), when consumers begin purchasing vaccine, $R_\infty(Q_c^*)$ reverses course and begins to slope downward in $\mathcal{R}_0$. The direct, epidemiological effect of an increase in $\mathcal{R}_0$ continues to be to increase $R_\infty(Q_c^*)$. Working in the oppositive direction is an indirect, economic effect of an increase in $\mathcal{R}_0$, inducing consumers to purchase more vaccine. In (SR2), the indirect effect is strong enough to dominate the direct effect, so that $R_\infty(Q_c^*)$ decreases in $\mathcal{R}_0$. Mathematically, to maintain the constant marginal private benefit ($MPB_c^* = c$) observed throughout (SR2), the increase in infectiveness $\mathcal{R}_0$ must be offset by a reduction in infections to maintain a constant probability of contracting the disease. In cases (SR3) and (SR4), $R_\infty(Q_c^*)$ again rises with $\mathcal{R}_0$ because the direct effect of an increase in infectiveness cannot be offset by an increase in Q_c^* given that all susceptibles are vaccinated. The marginal externality MEX_c^* exhibits an even more complex nonmonotonic pattern. The interplay between increasing infectiveness and increasing vaccine quantity generates two local maxima in the figure, with the global maximum occuring at the boundary between cases (SR1) and (SR2).

The next proposition summarizes the comparative-static effects of an increase in $\mathcal{R}_0$ on the steady-state equilibrium under perfect competition, showing that the observations from Figure 1 are quite general. Appendix A provides proofs for results not obvious from Table 1.[9]

Proposition 1. *Consider the comparative-static effect of $\mathcal{R}_0$ on equilibrium under perfect competition in the short-run analysis.*

- *Price and industry profit are constant, with $P_c^* = c$ and $\Pi_c^* = 0$.*
- *Q^c and MPB^c are weakly increasing in $\mathcal{R}_0$.*
- *$R_\infty(Q_c^*)$, MSB_c^*, and MEX_c^* are nonmonotonic in $\mathcal{R}_0$. $R_\infty(Q_c^*)$ attains a single interior local maximum, which is a global maximum if and only if $\tilde{c} \geq 1-\theta$. Each of MSB_c^* and MEX_c^* attain no more than two interior local maxima, one of which is a global maximum.*
- *W_c^* is weakly decreasing in $\mathcal{R}_0$.*

[9]The proofs for the nonmonotonic variables are particularly intricate since they involve characterizing complex shapes of functions without closed-form solutions. Our approach uses the concavity of the sign of the function's derivative to count the derivative's roots, coupled with an examination of the limiting value of the derivative.

For each Q_c^*, MPB_c^*, and W_c^*, there exists a nonempty interval of $\mathcal{R}_0$ such that the weak change is strict.

3.2. Monopoly

Since a monopolist charges a markup above cost, $P_m^* \geq c = P_c^*$, implying $Q_m^* \leq Q_c^*$. Thus, in case (SR1) in which $Q_c^* = 0$, we have $Q_m^* = 0$. Case (SR1) is thus trivially identical across perfect competition and monopoly. In the remaining cases, competitive firms are able to make positive sales at price c. By continuity, the monopolist can make positive sales at some small markup above c, implying $Q_m^* > 0$ for $\mathcal{R}_0$ in cases (SR2) and above.

To solve for Q_m^* in these other cases, we convert the monopolist's maximization problem so that the choice variable is quantity rather than price. The monopolist optimally sets a price to extract the entire private benefit of the marginal consumer, leading to inverse demand $P(Q) = MPB(Q)$. The monopolist chooses Q to maximize $[P(Q)-c]Q = [MPB(Q)-c]Q$ subject to $Q \leq \hat{S}_0$, a constrained maximization problem which can be solved using the Kuhn-Tucker method.

Differentiating the profit function with respect to Q yields

$$\theta H\left[1-\frac{S_\infty(Q)}{S_0(Q)}\right]\left\{1-\frac{\theta Q\mathcal{R}_0 S_\infty(Q)}{S_0(Q)[1-\mathcal{R}_0 S_\infty(Q)]}\right\}-c. \tag{26}$$

If the constraint $Q \leq \hat{S}_0$ does not bind, the monopoly quantity can be found by setting (26) equal to zero, which after manipulation yields

$$MPB(Q_m^*) = P(Q_m^*) = c \Big/ \left\{1-\frac{\theta Q_m^*\mathcal{R}_0 S_\infty(Q_m^*)}{S_0(Q_m^*)[1-\mathcal{R}_0 S_\infty(Q_m^*)]}\right\}. \tag{27}$$

Several useful insights are immediate consequences of (27). Since the denominator in braces is less than 1, we have $P_m^* = P(Q_m^*) > c = P_c^*$. If Q_c^* is an interior solution—i.e., $Q_c^* \in (0, \hat{S}_0)$, which is true in case (SR2) of Table 1—then the price relationship has a direct implication for quantities, namely, $Q_m^* < Q_c^*$.

Equation (27) can also be used to establish the existence of parameters for which the first best is attained under perfect competition but monopoly falls short. To see this, consider $\mathcal{R}_0$ on the boundary between cases (SR2) and (SR3) in Table B1, i.e., $\mathcal{R}_0 = |\ln(1-\tilde{c})|/[\hat{I}_0 + (1-\theta)\hat{S}_0]$. The entries for MPB_c^* in Table B1 imply $P(\hat{S}_0) = c$ at this value of $\mathcal{R}_0$, but (27) implies $P(Q_m^*) > c$. Hence, $P(Q_m^*) > P(\hat{S}_0)$, implying $Q_m^* < \hat{S}_0$ at this boundary value of $\mathcal{R}_0$. By continuity, there is

a neighborhood of $\mathcal{R}_0$ above this boundary value for which the first best is not obtained under monopoly but is obtained under perfect competition.

If the monopoly quantity is not an interior solution, then the constraint holds with equality, implying $Q_m^* = \hat{S}_0$. According the the Kuhn-Tucker conditions, the constraint holds with equality when the parameters are such that the derivative (26) is nonnegative when evaluated at the constraint quantity. Substituting $Q = \hat{S}_0$ into (26) and rearranging, this condition is

$$\Phi_I(\hat{S}_0)\left[1-\left(\frac{\theta}{1-\theta}\right)\frac{\mathcal{R}_0 S_\infty(\hat{S}_0)}{1-\mathcal{R}_0 S_\infty(\hat{S}_0)}\right] \geq \tilde{c}. \tag{28}$$

The left-hand side equals 1 in the limit $\mathcal{R}_0 \uparrow \infty$ by Lemma 8, exceeding the right-hand side since $\tilde{c} < 1$ by assumption (20). Therefore, for sufficiently high $\mathcal{R}_0$, a monopoly produces the first-best quantity, entailing universal vaccination of all susceptibles: $Q_m^* = \hat{S}_0 = Q^{**}$.

The values of equilibrium variables resulting from the preceding analysis are recorded in Table 2. The results reported in the column for cases (SR2) and (SR3) are not provided in analytic form, let alone in closed form. This need not preclude definitive comparative-statics results; one could apply the Implicit Function Theorem to the first-order condition (26) to determine how Q_m^* changes with $\mathcal{R}_0$ in (SR2) and (SR3). However, this approach does not deliver a definitive sign. We can be sure that Q_m^* increases in $\mathcal{R}_0$ for some $\mathcal{R}_0$ in (SR2) and (SR3)—since Q_m^* must rise from 0 to the first-best quantity $\hat{S}_0$ somewhere in that set by continuity—but we cannot rule out the possibility that the monopolist responds to an increase in $\mathcal{R}_0$ in some subintervals by reducing output in order to extract an even larger price increase than otherwise.

Despite these challenges, we are able to derive definitive comparative-statics results for Π_m^* by the envelope theorem. In addition, we are able to show that the nonmonotonic behavior of $R_\infty(Q^*)$, MSB^*, and MEX^* extends from perfect competition to monopoly. We report these comparative-statics results in the next proposition, proved in Appendix A.

Proposition 2. *Consider the comparative-static effect of $\mathcal{R}_0$ on equilibrium under monopoly in the short-run analysis.*

- $\Pi_m^* = 0$ *for* $\mathcal{R}_0 \leq |\ln(1-\tilde{c})|/(\hat{I}_0 + \tilde{c}\hat{S}_0)$*;* Π_m^* *is positive and strictly increasing for higher* $\mathcal{R}_0$*.*
- *If* $\tilde{c} \geq 1-\theta$*,* $R_\infty(Q_m^*)$ *is nonmonotonic in* $\mathcal{R}_0$ *and attains an interior global maximum.*
- MSB_m^* *and* MEX_m^* *are nonmonotonic, attaining interior global maxima.*

Equilibrium values of selected variables under monopoly are graphed as functions of $\mathcal{R}_0$ as the solid curves in Figure 1. The two market structures overlap in case (SR1), neither generating any vaccine output. The two market structures overlap again in (SR4), both generating the first-best quantity $Q^{**} = \mu$. In between—in (SR2) and (SR3)—the two market structures diverge, with monopoly generating strictly lower output, entailing more total infections over the epidemic, higher marginal private benefit, and lower welfare. The marginal externality can be considerably higher for monopoly for some $\mathcal{R}_0$ but can be slightly lower for some $\mathcal{R}_0$ as the lower monopoly output generates higher marginal private benefit, leaving less residual externality.

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 monopolist's exercise of this market power generates deadweight loss. The large gap between W_c^* and W_m^* for an intermediate range of $\mathcal{R}_0$ in the bottom panel of Figure 1 suggests that monopoly distortions may be worst for moderate levels of infectiveness. Market power generates little welfare loss for either the lowest or highest values of $\mathcal{R}_0$.

The graph of W^* under monopoly illustrates the remarkable possibility that increasing $\mathcal{R}_0$ can increase welfare. One would think that society would always be harmed by an increase in infectiveness. While the direct, epidemiological effect of an increase in $\mathcal{R}_0$ harms society, the indirect effect of increasing vaccinations can counteract the direct effect, increasing welfare over some parameter ranges. In the bottom panel of Figure B1, we see this possibility emerging for higher values of $\mathcal{R}_0$ in (SR3). Under monopoly, not only do consumers fail to consider the external benefit their vaccination provides other consumers, but the monopolist 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 $\mathcal{R}_0$ can provide such a large indirect benefit that it swamps the direct harm from an increase in $\mathcal{R}_0$, leading to an increase in social welfare.

4. Government Subsidies

We have seen that the positive externality associated with vaccine consumption can lead to under-consumption relative to the first best under both perfect competition and monopoly. This naturally raises the question of whether the government can intervene to correct the market failure. In this section, we characterize the optimal government subsidy and determine its comparative-static properties.

Assume a benevolent government with the objective of maximizing social welfare commits to a per-dose subsidy $GS \geq 0$ at the outset of the game. Adopting the accounting convention that the subsidy is paid to firms, the subsidy is equivalent to a reduction in firms' marginal cost from c to $c - GS$. Since social welfare is maximized by the first-best quantity Q^{**}, the first-best subsidy GS^{**} is that implementing Q^{**}. To accommodate cases in which the government is indifferent among a possibly open set of subsidies maximizing social welfare, we take GS^{**} to be the infimum of the set.

It is straightforward to establish a set of broad results for any market structure. Since $MEX(Q) \geq 0$, equilibrium output Q^* cannot exceed the first best Q^{**}. If, in addition, $Q^{**} = 0$, then $0 \leq Q^* \leq Q^{**} = 0$, implying $Q^* = Q^{**}$, in turn implying $GS^{**} = 0$ since the first best can be achieved without a subsidy. The proof of the next proposition shows that $Q^{**} = 0$ for all $\mathcal{R}_0$ in a neighborhood above 0. For sufficiently small $\mathcal{R}_0$, then, $GS^{**} = 0$ for any market structure.

We can also draw broad conclusions about the optimal subsidy for high values of $\mathcal{R}_0$. Suppose monopoly output is the corner solution $Q_m^* = \hat{S}_0$. Then $\hat{S}_0 = Q_m^* \leq Q^{**} \leq \hat{S}_0$, implying $Q_m^* = Q^{**}$, in turn implying $GS^{**} = 0$ since the first best can be achieved without a subsidy under monopoly. By Table 2, monopoly attains the first best for all $\mathcal{R}_0$ in case (SR4), which the text argued includes an interval of sufficiently high values of $\mathcal{R}_0$. We conclude that for sufficiently high $\mathcal{R}_0$, $GS_m^{**} = 0$. This result also immediately extends to perfect competition or any market structure involving weakly higher output than monopoly.

Having established that $GS^{**} = 0$ for intervals of low and high values of $\mathcal{R}_0$ for general market structures, if it can be shown that $GS^{**} > 0$ for some intermediate value of $\mathcal{R}_0$, it is immediate that GS^{**} is nonmonotonic, attaining a maximum for some interior $\mathcal{R}_0 \in (0, \infty)$ as the next proposition states. The proof provided in Appendix A fills in this and other omitted details.

Proposition 3. *For monopoly—or any market structure involving weakly lower output including perfect competition—GS^{**} is nonmonotonic in $\mathcal{R}_0$, equaling 0 for sufficiently low and sufficiently high $\mathcal{R}_0$, and attaining a positive maximum for some $\mathcal{R}_0 \in (0, \infty)$.*

As the proposition indicates, the optimal subsidy is not monotonically increasing in $\mathcal{R}_0$ as might be inferred based solely on epidemiological considerations but is hump shaped. The difficulty in addressing a disease depends not only on its infectiousness but also on consumers' response to this infectiousness. Consumers respond to extremely infectious diseases by getting vaccinated even if many others also do. Moderately infectious diseases provide consumers more leeway to free ride on the vaccination of others.

Turn to a more precise characterization of the optimal subsidy under perfect competition. Suppose the first best is given by an interior solution, i.e., $Q^{**} \in (0, \hat{S}_0)$. Then Q^{**} must satisfy the first-order condition for welfare maximization $MSB(Q^{**}) = c$, implying $MPB(Q^{**}) + MEX(Q^{**}) = c$, in turn implying $P_c^{**} = MPB(Q^{**}) = c - MEX(Q^{**})$. Since competitive firms pass the subsidy through to consumers, $P_c^{**} = c - GS_c^{**}$. Combining the preceding equations yields $GS_c^{**} = MEX(Q^{**})$, the familiar result that setting the subsidy equal to the marginal externality is optimal. If Q^{**} is a corner but Q_c^* is not, i.e., $Q_c^* < \hat{S}_0 = Q^{**}$, then the highest price at which output $\hat{S}_0$ is purchased satisfies $P_c^{**} = MPB(\hat{S}_0)$. Combined with competitive pass through, $P_c^{**} = c - GS_c^{**}$, we have $GS_c^{**} = c - MPB(\hat{S}_0)$. If $Q_c^* = Q^{**} = \hat{S}_0$, then the preceding proposition implies $GS_c^{**} = 0$. The various results for $Q^{**} = \hat{S}_0$ can be nested as $GS_c^{**} = \max[0, c - MPB(\hat{S}_0)]$.

Next, turn to a more precise characterization of the optimal subsidy under monopoly. The monopolist regards the subsidy as a reduction in marginal cost, maximizing $[P(Q) - c + GS]Q = [MPB(Q) - c + GS]Q$. To generate the first best, the optimal subsidy GS_m^{**} must force the monopoly's first-order condition to be satisfied by Q^{**}:

$$MPB(Q^{**}) - c + GS_m^{**} + Q^{**} \frac{\partial MPB(Q^{**})}{\partial Q} = 0. \tag{29}$$

If $Q^{**} \in (0, \hat{S}_0)$, then the analysis of perfect competition showed $MPB(Q^{**}) = P_c^{**} = c - GS_c^{**} = c - MEX(Q^{**})$. Substituting into (29) yields, after rearranging,

$$GS_m^{**} = MEX(Q^{**}) \left(\frac{\hat{S}_0}{\hat{S}_0 - \theta Q^{**}} \right). \tag{30}$$

Equation (30) also uses the fact that $\partial MPB(Q)/\partial Q = -\theta MEX(Q)/S_0(Q)$ for all Q, as can be seen by combining (17) and (24). Equation (30) shows that the monopoly subsidy is proportional to the marginal externality, scaled up by the factor in parentheses, which adjusts for the monopoly markup. This scale factor grows without bound as the first best approaches successful vaccination of all initial

susceptibles. Equation (30) also implies that $GS_m^{**} > GS_c^{**}$ since $GS_c^{**} = MEX(Q^{**})$.

If $Q^{**} = \hat{S}_0$, then GS_m^{**} must force (29) to hold at that corner value of quantity. The next proposition summarizes the analysis for the two market structures.

Proposition 4. *If $Q^{**} = 0$, then $GS_c^{**} = GS_m^{**} = 0$. If $Q^{**} \in (0, \hat{S}_0)$, then $GS_c^{**} = MEX(Q^{**})$ and $GS_m^{**} = MEX(Q^{**})\hat{S}_0/(\hat{S}^0 - \theta Q^{**})$. If $Q^{**} = \hat{S}_0$, then $GS_c^{**} = \max[0, c - MPB(\hat{S}_0)]$ and*

$$GS_m^{**} = \max\left[0, c - MPB(\hat{S}_0) + \left(\frac{\theta}{1-\theta}\right) MEX(\hat{S}_0)\right]. \tag{31}$$

Across all cases, $GS_m^{**} \geq GS_c^{**}$ with strict inequality if $Q^{**} > 0$ and $Q_m^* < \hat{S}_0$.

5. Increasing Social Returns

Typical products exhibit concave social benefits. The underlying logic is that initial units provide higher marginal social benefits than subsequent units since highest-value uses are served first, with subsequent units allocated to lower-value uses. Epidemiological externalities may lead this logic to fail with vaccines. Vaccinating a few individuals may do little to slow the spread of an epidemic if susceptibles are likely contract the disease from the many remaining unvaccinated people in any event. Doubling coverage may more than double the social benefit if the additional coverage is needed to make a dent in the infection rate.

In this section, we analyze conditions under which vaccines exhibit increasing rather than diminishing social returns. To this point we have assumed that any amount of vaccine can be produced at the constant marginal cost c. In reality, capacity constraints may prevent production up to the point that marginal social benefit equals production cost; rationing may be required. With the population divided into regional subunits experiencing relatively independent epidemiological processes because of restricted travel flows, it is natural to ask whether vaccine should be spread across regions in proportion to their populations (as considerations of fairness or heterogeneity in value within each region might dictate) or whether the benefits would be larger if vaccine were concentrated in fewer regions (chosen by lottery if urgency of need in certain regions does not provide sufficient reason for concentrating vaccine there).

Formally, a vaccine exhibits increasing social returns if $MSB(Q)$ is increasing in Q. Differentiating (23), substituting from (14), and rearranging yields

$$\frac{\partial MSB(Q)}{\partial Q} = \frac{\theta^2 H \mathcal{R}_0 S_\infty(Q) S_0(Q) \Phi_I(Q)}{[1 - \mathcal{R}_0 S_\infty(Q)]^3} \{\mathcal{R}_0[S_0(Q) + S_\infty(Q)] - 2\}. \tag{32}$$

By Lemma 5, all the factors on the right-hand side are definitively positive except for the last. Thus, the sign of the last factor in braces determines whether the vaccine exhibits increasing social returns. Rearranging gives the following proposition.

Proposition 5. *The Qth unit of vaccine exhibits increasing social returns if and only if*

$$\mathcal{R}_0 \left[\frac{S_0(Q) + S_\infty(Q)}{2} \right] > 1. \tag{33}$$

Earlier, we identified the inequality $\mathcal{R}_0 S_0(Q) > 1$ as the condition for epidemic expansion, requiring a combination of sufficient infectiveness, $\mathcal{R}_0$, and sufficient fuel, $S_0(Q)$, for the disease to spread. Condition (33) is more stringent, requiring that the product of $\mathcal{R}_0$—not with the initial number of susceptibles—but with the average of the initial and eventual number of susceptibles exceed 1. One can see this is more stringent since $S_\infty(Q) < S_0(Q)$ by Lemma 4. Proposition 5 can thus be interpreted as saying that unit Q of the vaccine exhibits increasing social returns if the potential for epidemic expansion is sufficiently high at the current output level.

The next proposition provides simpler sufficient conditions for the vaccine to exhibit increasing social returns at initial output levels and at all output levels. It is proved in Appendix A as a straightforward corollary of Proposition 5.

Proposition 6. *The vaccine exhibits initial increasing social returns (i.e., at an output level of $Q = 0$) if $\mathcal{R}_0 \hat{S}_0 \geq 2$. The vaccine exhibits everywhere increasing social returns (i.e., at all output levels $Q \in (0, \hat{S}_0)$) if $\mathcal{R}_0 \hat{S}_0 \geq 2/(1-\theta)$.*

According to Proposition 6, if a federal authority only has access to a small stockpile of a vaccine to allocate across several similar states with independent epidemiological processes, allocating the entire stockpile to one state would produce more social benefit than spreading it evenly across them if $\mathcal{R}_0 \hat{S}_0 > 2$. If, for example, $\hat{S}_0 = 0.8$ in each state, then concentrating the vaccine would be efficient for any $\mathcal{R}_0 > 2.5$. If the more stringent condition $\mathcal{R}_0 \hat{S}_0 > 2/(1-\theta)$ holds, then even a starker form of concentration is efficient: not just for very small stockpiles but for any size, the federal authorities should vaccinate all susceptibles in one state before moving to the next. The starkness of the policy hinges on the modeled consumer homogeneity: if each state has some vulnerable consumers with a high benefit from vaccinating, a higher bar on $\mathcal{R}_0$ would need to be cleared for concentrating vaccines in one state to be more efficient than serving high-value consumers everywhere first.

Proposition 6 hints that efficacy θ plays a role in determining whether social returns are everywhere increasing. The next proposition, also proved in Appendix A as a straightforward corollary

of Proposition 5, draws a clearer connection.

Proposition 7. *A perfectly effective vaccine ($\theta = 1$) cannot exhibit everywhere increasing social returns.*

The result is intuitive. Protecting all but the last consumer with a perfectly effective vaccine eliminates the externality from vaccinating the marginal consumer, the force behind increasing social returns. Thus, social returns must start to diminish with a perfectly effective vaccine at some point before universal vaccination is reached.

Proposition 7 does not automatically imply that a perfectly effective vaccine should be spread across regions. Marginal social benefits may be diminishing in some geographic region yet still be above that in other regions not yet receiving a vaccine.

6. Vaccines Versus Drugs

Commentators on the pharmaceutical industry frequently suggest that firms are biased in favor of developing drugs rather than vaccines. Kremer and Snyder (2015, 2018) list a variety of reasons for this bias, ranging from vaccines' complexity relative to drug molecules, to the scale often needed for vaccine clinical trials, to the evaporation of consumers' private disease-risk information when making drug purchases (the focus of those papers).

The epidemiological externality analyzed in this paper provides another rationale. By preventing individuals from becoming infected, vaccines curtail their transmission of the disease to others. The reduction in others' disease risk is a public good that reduces others' willingness to pay for a vaccine. This public-good feature distinguishes vaccines from some drugs that treat symptoms without curing the underlying disease or inhibiting transmission. Firms would have more of incentive to develop a drug that does not have this demand-reducing public-good feature than a similarly effective vaccine.[10]

To quantify a monopoly's bias toward a drug and against a vaccine, consider a drug that is similar in all ways to the vaccine analyzed to this point except that the drug does not reduce disease transmission. Finding the right normalization to make drug and vaccine costs equivalent is somewhat

[10]If one takes the focus on steady states in the long-run analysis as literally implying that the discount rate is zero, then there would not be any bias in development decisions. All products producing any positive flow profit would be developed regardless of how large is 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.

delicate since at equal marginal production costs c the total cost of serving a population with a drug is lower if it only needs to be administered to infected consumers rather than the whole population in advance as with a vaccine. We finesse this normalization issue by assuming both products are costless to produce and administer, i.e., $c = 0$. Assume the drug is effective with probability θ. 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.

Having computed monopoly profit and welfare from a vaccine, respectively Π^*_{mv} and W^*_{mv}, it remains to compute the analogous variables for a drug, respectively Π^*_{md} and W^*_{md}. For all $\mathcal{R}_0 > 0$, the drug monopoly can charge $P^*_{md} = \theta H$ to the $\hat{I}_0$ individuals infected at the moment the drug is developed as well as the $\hat{S}_0 - S_\infty(0)$ individuals who become infected at some point afterwards, yielding drug profit

$$\Pi^*_{md} = \theta H \left[\hat{I}_0 + \hat{S}_0 - S_\infty(0) \right]. \tag{34}$$

To compute equilibrium welfare with a drug, the $\hat{I}_0$ individuals infected initially along with the $\hat{S}_0 - S_\infty(0)$ infected later obtain health benefit H with probability θ from the drug. The $S_\infty(0)$ remaining susceptibles are never infected and obtain health benefit H with certainty, yielding the following expression for equilibrium welfare after rearranging:

$$W^*_{md} = H \left[(1-\theta) S_\infty(0) + \theta(\hat{I}_0 + \hat{S}_0) \right]. \tag{35}$$

Comparing these expressions against the analogous entries in Table 2 for a vaccine leads to the next proposition. Details behind the proof are provided in Appendix A. The proposition uses the notation $\Delta\Pi^*_m = \Pi^*_{md} - \Pi^*_{mv}$ and $\Delta W^*_m = W^*_{md} - W^*_{mv}$ for differences between equilibrium variables for the two products and $\Delta W^{**} = W^{**}_d - W^{**}_v$ for difference between first-best welfare.

Proposition 8. *Suppose $c = 0$. For all $\mathcal{R}_0 > 0$, $\Delta\Pi^*_m > 0$. $\Delta\Pi^*_m$ is nonmonotonic in $\mathcal{R}_0$, with $\lim_{\mathcal{R}_0 \downarrow 0} \Delta\Pi^*_m = \lim_{\mathcal{R}_0 \uparrow \infty} \Delta\Pi^*_m = \inf_{\mathcal{R}_0 > 0} \Delta\Pi^*_m = \theta H \hat{I}_0$. For extreme values of $\mathcal{R}_0$, ΔW^*_m is positive: $\mathcal{R}_0$, $\lim_{\mathcal{R}_0 \downarrow 0} \Delta W^*_m > 0$ and $\lim_{\mathcal{R}_0 \uparrow \infty} \Delta W^*_m > 0$. However, there exist parameters for which $\Delta W^*_m < 0$.*

According to the proposition, the monopoly is biased toward the drug for all parameters, and this bias leads the firm to choose the socially inferior product for some parameters. For other parameters, the drug provides higher welfare than the vaccine. Two such cases are provided by the extremes $\mathcal{R}_0 \downarrow 0$ and $\mathcal{R}_0 \uparrow \infty$, examined in turn. Equilibrium welfare never falls below $\theta H \hat{I}_0$ for a

drug monopoly, even for extreme values of $\mathcal{R}_0$. Administering a drug to the $\hat{I}_0$ initially infected provides a social benefit even if $\mathcal{R}_0$ is so low that the infection does not spread to others. A vaccine cannot provide this social benefit because it is useless unless administered prior to infection in the model. Thus, equilibrium welfare is higher with a drug than vaccine in the limit $\mathcal{R}_0 \downarrow 0$. Equilibrium welfare is also higher with a drug than vaccine in the limit $\mathcal{R}_0 \uparrow \infty$. The externality associated with vaccine disappears because susceptibles are certain to contract the disease, if no one else, from an unsuccessfully vaccinated person with an infinitely infective disease. Hence, apart from the drug's remaining social benefit of treating the $\hat{I}_0$ initially infected, the drug and vaccine provide equal welfare in the limit $\mathcal{R}_0 \uparrow \infty$. The opposing welfare factors—the drug helps initially infected but the vaccine reduces subsequent spread to others—prevent many firm conclusions from being drawn about the sign of the equilibrium or first-best welfare differentials.

7. Covid Calibration

This section provides an illustrative calibration using parameters drawn from the current Covid pandemic. The calibration is meant more as illustration than a forecast. Our present model is too stylized on many fronts to provide accurate forecasts, abstracting 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.

We take estimates of needed parameters as of October 2020, calibrating the counterfactual effect of the arrival of a vaccine when emergency use was starting to be approved for the available Covid vaccines. We use estimates from U.K. government agencies, which provide some of the best estimates for a developed country then available. Based on U.K. Government Office for Science (2020), we set $\mathcal{R}_0 = 1.5$. Based on U.K. Office for National Statistics (2020), we take the proportion of infected at that time to be $\hat{I}_0 = 0.19\%$ and the proportion of recovered to be $\hat{R}_0 = 6.2\%$, implying $\hat{S}_0 = 1 - \hat{I}_0 - \hat{R}_0 = 93.6\%$. Based on Public Health England (2021), we set $\theta = 0.8$, the midpoint of the range of estimated efficacy of two doses of the Pfizer vaccine against Covid infection (including both symptomatic and asymptomatic), tested when the Alpha variant dominated. 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.[11]

One can show that these parameters lead to case (SR3) of Tables 1 and 2. In this case, the first best is obtained under perfect competition. Universal vaccination results in all current susceptibles being protected through the end of the pandemic. The first best is not obtained under monopoly. Using numerical methods to compute $S(Q)$ in (13) and to optimize monopoly profit, we find that the monopoly price is set to 23% of the harm from contracting the disease. At this price, only 21% of susceptible consumers buy, generating welfare equal to 71% of the available health benefit. The optimal subsidy required to generate the first best under monopoly is 55% of the equilibrium monopoly price.

Since $\mathcal{R}_0\hat{S}_0 = 1.4$, the simple necessary condition for initially increasing social returns given in Proposition 6, $\mathcal{R}_0\hat{S}_0 > 2$, is not met. However, the more complicated but weaker condition (33) provided by Proposition 5 does indicate initially increasing returns, as $\mathcal{R}_0[\hat{S}_0 + S_\infty(0)]/2 = 1.04 > 1$. Examing (33) for a range of quantities shows that increasing returns persists through output equal to 22% of the susceptible population. Supposing that a stockpile has to be allocated to two identical states with independent epidemiological processes, concentrating the entire stockpile in one state generates higher welfare than dividing equally until the stockpile exceeds 31% of the population of one state. Larger stockpiles than this are more efficiently divided equally between the states. The results do not suggest Covid vaccines have strongly increasing social returns.

For the calibrated parameters, $\Delta\Pi_m^* > 0$, so a monopoly firm would be biased toward a drug and away from a vaccine. Given and $\Delta W_m^* > 0$, welfare is higher with a drug, so the bias does not lead to a distortion.

8. Conclusion

We analyzed the market for technologies preventing individuals from contracting a disease. Such products are interesting since, by preventing the consumer from contracting a disease, they exert

[11] Castillo et al. (2021) report that prices for available Covid vaccines were no greater than \$40 per course. Health losses can be computed following Snyder, *et al*. (2020). Hanlon *et al*. (2021) estimate 12 years of lost life (YLL) per death. Since this estimate already allocates shorter lifespans to people with comorbidities, we assume one YLL translates into one disability adjusted life year (DALY) without need for further downward adjustment to reflect a proportion of years lived with a disability. To convert DALYs into monetary values, we multiply DALYs lost in a country by three times that country's 2019 GDP per capita, reflecting World Health Organization (WHO) standards for a cost-effective health intervention in a country stated in Marseille *et al*. (2015). According to this standard, a health intervention is cost effective if the cost per DALY saved is less than three times that country's per-capita GDP (\$65,253 in the U.S. in 2019). Putting these estimates together yields an estimate of $H = 12 \times 3 \times \$65,253 = \$2.35$ million. Using the calibrated value of $\theta = 0.8$ yields $\tilde{c} = 40/(0.8 \times 2.35 \times 10^6) = 2.13 \times 10^{-5}$.

a positive externality, reducing the spread to others. Though the analysis applies to a variety of technologies such as circumcision, bed nets, or social distancing, the discussion focused on vaccines for concreteness. Vaccines (and most of the other aforementioned technologies) are not pure public goods since they are physical products that exhibit rivalry and excludability in consumption, yet they share with public goods the feature that one's consumption reduces others' demand for that product, a feature that can potentially lead to large distortions in firms' supply decisions.

Such distortions and policy correctives were the focus of this paper. To study them, we constructed and analyzed a theoretical model of the vaccine market involving economic agents that base their consumption and production decisions on rational expectations of the disease's evolution consistent with a standard SIR epidemiological model. We provided two separate analyses. Our short-run analysis studied a vaccine campaign launched at one point in time along the path of an epidemic disease such as Covid-19 that is expected to wane before population turnover becomes a relevant issue. Our long-run analysis studied of the steady-state equilibrium for an endemic disease such as HIV. We sought to provide a comprehensive account of equilibrium variables such as price, quantity, profit, and welfare across a variety of market structures ranging from perfect competition to Cournot to monopoly and to study how those variables changed in response to parameter changes.

A key variable was the equilibrium marginal externality. We consistently found—across both short- and long-run analyses, across a range of market structures, and across models with either homogeneous or heterogeneous consumers—that the equilibrium marginal externality is nonmonotonic in the infectiousness of the disease as measured by $\mathcal{R}_0$. For low levels of $\mathcal{R}_0$, one consumer's vaccination provides little benefit to others because there is little chance the consumer would have infected them anyway. For high levels of $\mathcal{R}_0$, one consumer's vaccination provides little benefit to others because they will most likely contract it from a different source anyway. The marginal externality is greatest for intermediate values of $\mathcal{R}_0$. This nonmonotonicity carries over to other outcome variables such as GS^{**} (the minimal subsidy necessary to obtain the first-best vaccine quantity) and $\Delta\Pi_m^*$ (the difference between monopoly profit from a drug that does not exert the epidemiological externality and from a vaccine that does). Diseases with moderate infectiousness may exhibit the greatest distortions and be prime targets for subsidy.

If the negotiated price reflects the threat point of decentralized vaccine sales to individuals on the private market, our results on the relative profitability of vaccines versus drugs gain relevance. The positive epidemiological externality can lead a vaccine to be less lucrative than a drug lacking that externality. The monopoly's threat point may be worse with a vaccine, implying that the hold-

up problem and consequent underinvestment may be worse with a vaccine than a drug. To address the hold-up problem, governments may have to consider negotiating with vaccine manufacturers ex ante, prior to substantial investment in R&D and production capacity.

While the long- and short-run analyses share many key qualitative findings, there are points of contrast. Many of these points of contrast can be traced to the presence of the $\hat{I}_0$ infected individuals had yet recovered at vaccine rollout and thus capable of transmitting disease to unvaccinated. This group is irrelevant in the long-run analysis because recovery and population turnover removes them from the steady state. The presence of this group leads to the possibility that universal vaccination of susceptibles with a perfectly effective vaccine can be a viable business strategy since not all disease reserviors are eliminated. Indeed, Table B2 shows the monopoly equilibrium involves universal vaccination for sufficiently high $\mathcal{R}_0$. The presence of this group raises the possibility that a vaccine with a positive epidemiological externality can be welfare-dominated by a drug without it: if the externality is small, welfare may be driven by the advantage of the drug in treating the $\hat{I}_0$ infecteds for whom the vaccine arrives too late to help, assuming the vaccine must be administered prior to infection to be effective.

The short-run analysis provided fertile ground for understanding when vaccination exhibits increasing social returns. According to Proposition 5, a vaccine exhibits initially increasing social returns if $\mathcal{R}_0\hat{S}_0 \geq 2$ and everywhere increasing social returns if $\mathcal{R}_0\hat{S}_0 \geq 2/(1-\theta)$. If the first condition holds, a small capacity should be concentrated in a single region; and if the second condition holds, a first region should be completely served before moving to a second regardless of capacity size. These stark implications for concentrating supplies hinge on the homogeneity of consumers in the model but raise the possibility o equitable allocation leading to inefficiency.

In our calibration to a Covid vaccine, we found that the first best would be approached by a competitively supplied vaccine but a monopoly (selling to individuals on the private market, not bulk sales to governments) would set such a high price that only 21% of susceptibles would buy. A vaccine would have increasing social returns not everywhere but through 31% of the susceptible population. The presence of increasing social returns argues for subsidizing aggressive investment to boost capacity beyond this point if concentrating supplies in few countries is either unpalatable or outweighed in by the benefit of vaccinating vulnerable subpopulations in every country.

In calibrations to endemic diseases, we found that strong cases can be made for subsidizing competitively supplied technologies such as the example of circumcisions to prevent HIV. However, with a monopoly supplier, the per-unit subsidy needed to obtain the first best can be prohibitively

expensive. For example, in a calibration of the long-run analysis to measles, we found that the subsidy required to obtain the first best would be fifteen times the harm from certainly contracting the disease. In the short-run analysis, Section 4 provided a general argument that in the limit with a perfectly effective vaccine, if a monopolist would not produce the first best in equilibrium without government intervention, it cannot be induced to produce the first best by any finite per-unit subsidy. With even a small deadweight loss of taxation, such large subsidies would fail a cost-benefit test. Our analysis thus suggests that, in many cases, governments may either need to give up on the first best as a feasible target for a per-unit subsidy or need to consider other policies such as purchasing vaccines in bulk at a negotiated price building in a lump-sum subsidy.

Appendix A. Proofs

To avoid an excess of technical detail in the text, the formal proofs of most lemmas and propositions, as well as verification of Table 1, have been collected in this appendix.

Proof of Lemma 1

We begin by proving the claims about $I_t(Q)$. Substituting (11) into (3) yields

$$\dot{I}_t(Q) = \frac{1}{\alpha} I_t(Q)[\mathcal{R}_0 S_t(Q) - 1], \tag{A1}$$

or, rearranging,

$$\frac{\dot{I}_t(Q)}{I_t(Q)} = \alpha[\mathcal{R}_0 S_t(Q) - 1]. \tag{A2}$$

Recognizing the left-hand side as $\partial \ln I_t(Q)/\partial t$ and integrating yields

$$\int_0^t \frac{\partial \ln I_\tau(Q)}{\partial \tau} d\tau = \int_0^t \alpha[\mathcal{R}_0 S_\tau(Q) - 1] d\tau. \tag{A3}$$

Invoking the Fundamental Theorem of Calculus, taking exponentials, and rearranging yields, for all $t \geq 0$,

$$I_t(Q) = I_0(Q) \exp\left(\int_0^t \alpha[\mathcal{R}_0 S_\tau(Q) - 1] d\tau \right). \tag{A4}$$

Since $I_0(Q) = \hat{I}_0 > 0$ by assumption, $I_t(Q)$ is the product of two positive factors.

Turn next to proving the claims about $S_t(Q)$. Rearranging (4), $I_t(Q) = \dot{R}_t(Q)/\alpha$. Substituting into (2) and rearranging yields $\dot{S}_t(Q)/S_t(Q) = -(\beta/\alpha)\dot{R}_t(Q) = -\mathcal{R}_0 \dot{R}_t(Q)$ by (11). Recognizing $\dot{S}_t(Q)/S_t(Q) = \partial \ln S_t(Q)/\partial t$ and integrating between $t' \geq 0$ and $t'' \geq t'$ yields

$$\int_{t'}^{t''} \frac{\partial \ln S_\tau(Q)}{\partial \tau} d\tau = -\int_{t'}^{t''} \frac{1}{\alpha} \dot{R}_\tau(Q) d\tau. \tag{A5}$$

Invoking the Fundamental Theorem of Calculus, taking exponentials, and rearranging yields

$$S_{t''}(Q) = S_{t'}(Q) e^{\mathcal{R}_0[R_{t'}(Q) - R_{t''}(Q)]}. \tag{A6}$$

Substituting $t' = 0$ and $t'' = t$ into (A6) yields

$$S_t(Q) = S_0(Q) e^{\mathcal{R}_0[R_0(Q) - R_t(Q)]}. \tag{A7}$$

Now $S_0(Q) = \hat{S}_0 - \theta Q > \hat{S}_0 - Q \geq 0$, where the first step holds by (6), the second by $\theta < 1$, and the third by $Q \in [0, \hat{S}_0]$. The right-hand side of (A7) is thus the product of two positive factors. *Q.E.D.*

Proof of Lemma 2

Substituting $I_t(Q) > 0$ into (4) yields $\dot{R}_t(Q) > 0$, implying $R_{t''}(Q) > R_{t'}(Q)$ for $t'' > t'$, implying $e^{\mathcal{R}_0[R_{t'}(Q) - R_{t''}(Q)]} < 1$. Since $S_{t'}(Q) > 0$ by Lemma 1, $S_{t''}(Q) \leq S_{t'}(Q)$ by (A6). *Q.E.D.*

Proof of Lemma 3

Since $I_t(Q) > 0$ by Lemma 1, by (A1), the sign of $\dot{I}_t(Q)$ is determined by the value of $\mathcal{R}_0 S_t(Q)$ relative to 1. First, suppose $\mathcal{R}_0 S_0(Q) \leq 1$. Consider any $t > 0$. Lemma 2 implies $S_t(Q) < S_0(Q)$, in turn implying $\mathcal{R}_0 S_t(Q) < \mathcal{R}_0 S_0(Q) \leq 1$. Substituting $\mathcal{R}_0 S_t(Q) < 1$ into (A1) implies $\dot{I}_t(Q) < 0$.

Next, suppose $\mathcal{R}_0 S_0(Q) > 1$. Substititing into (A1) for $t = 0$ implies $\dot{I}_0(Q) > 0$. By Martcheva (2015, p. 13), $I(Q) = 0$. Therefore, we must have $\dot{I}_t(Q) < 0$ for some $t > 0$. Thus, by continuinity, $\dot{I}_T(Q) = 0$ for some $T > 0$. Since $S_t(Q)$ is strictly decreasing by Lemma 1, T is unique, given by the value of t for which (A1) equals 0, implying T satisfies $S_T(Q) = 1/\mathcal{R}_0$. Since $S_t(Q)$ is strictly decreasing we have $\dot{I}_t(Q) > 0$ for all $t \in [0, T)$, $\dot{I}_t(Q) < 0$ for all $t > T$, and $I_T(Q)$ is the maximum infection rate. *Q.E.D.*

Proof of Lemma 4

See Martcheva (2015, p. 13) for a proof that $I_\infty(Q) = 0$. Martcheva (2015, p. 12) argues that the fact that $S_t(Q)$ is positive and montone implies that the limit $S_\infty(Q)$ exists.

To prove the remaining claim in the lemma, take the limit $t \uparrow \infty$ in (A7):

$$S_\infty(Q) = S_0(Q) e^{\mathcal{R}_0[R_0(Q) - R_\infty(Q)]}. \tag{A8}$$

By Lemma 1, $S_0(Q) > 0$. The proof of Lemma 2 showed that $R_t(Q)$ is strictly increasing in t. Thus $R_\infty(Q) > R_0(Q)$, implying $S_\infty(Q) < S_0(Q)$ by (A8). *Q.E.D.*

Proof of Lemma 5

First, suppose $S_0(Q) \leq 1/\mathcal{R}_0$. Then $S_\infty(Q) < 1/\mathcal{R}_0$ because $S_t(Q)$ is strictly decreasing in t by Lemma 2. Next, suppose $S_0(Q) > 1/\mathcal{R}_0$. In the last paragraph of the proof of Lemma 3, we proved the existence of $T > 0$ such that $S_T(Q) = 1/\mathcal{R}_0$. Since $S_t(Q)$ is strictly decreasing by Lemma 2, we have $S_\infty(Q) < S_T(Q) = 1/\mathcal{R}_0$. *Q.E.D.*

Proof of Lemma 6

Substituting (2) and (11) into (3) yields

$$\dot{I}_t(Q) = \frac{\dot{S}_t(Q)}{\mathcal{R}_0 S_t(Q)} - \dot{S}_t(Q). \tag{A9}$$

Integrating (A9) over $t \in [0, \infty)$ and applying the Fundamental Theorem of Calculus,

$$I_\infty(Q) - I_0(Q) = \frac{1}{\mathcal{R}_0}[\ln S_\infty(Q) - \ln S_0(Q)] - S_\infty(Q) + S_0(Q). \tag{A10}$$

Substituting $I_0(Q) = \hat{I}_0$ by (7), noting $I_\infty(Q) = 0$ by Lemma 4, and rearranging yields

$$\ln S_\infty(Q) - \mathcal{R}_0 S_\infty(Q) = \ln S_0(Q) - \mathcal{R}_0[\hat{I}_0 + S_0(Q)]. \tag{A11}$$

Further substituting $S_0(Q) = \hat{S}_0 - Q$ from (6) yields (12).

To derive (13), exponentiating both sides of (A11) and rearranging yields

$$S_\infty(Q) = \left\{ S_0(Q) e^{-\mathcal{R}_0[\hat{I}_0 + S_0(Q)]} \right\} e^{\mathcal{R}_0 S_\infty(Q)}, \tag{A12}$$

or, equivalently,

$$x = be^{ax}, \tag{A13}$$

where $x = S_\infty(Q)$, $a = \mathcal{R}_0$, and $b = S_0(Q)e^{-\mathcal{R}_0[\hat{I}_0 + S_0(Q)]}$. It is well-known that (A13) has solution $x = -\bar{L}(-ab)/a = |\bar{L}(-ab)|/a$, where the second equality holds if $a, b > 0$ implying $\bar{L}(-ab) < 0$. Substituting for x, a, and b in this solution as well as $S_0(Q) = \hat{S}_0 - Q$ from (6) yields (13).

Equation (A13) also has a solution in terms of the lower branch of the Lambert W function, $x = -\underline{L}(-ab)/a$. We reject this solution because it exceeds 1, which is out of bounds for $S_\infty(Q)$. *Q.E.D.*

Proof of Lemma 8

We will prove $\lim_{\mathcal{R}_0 \uparrow \infty}[\mathcal{R}_0 S_\infty(Q)] = 0$. The claim that $\lim_{\mathcal{R}_0 \uparrow \infty} S_\infty(Q) = 0$ follows as a direct consequence (a separate proof was also proved in the text). We have

$$\lim_{\mathcal{R}_0 \uparrow \infty}[\mathcal{R}_0 S(\hat{S}_0)] = \lim_{\mathcal{R}_0 \uparrow \infty} \left| \bar{L}\left(-\mathcal{R}_0(1-\theta)\hat{S}_0 e^{-\mathcal{R}_0[\hat{I}_0 + (1-\theta)\hat{S}_0]} \right) \right| \tag{A14}$$

$$= \left| \bar{L}\left(-(1-\theta)\hat{S}_0 \lim_{\mathcal{R}_0 \uparrow \infty} \frac{\mathcal{R}_0}{e^{\mathcal{R}_0[\hat{I}_0 + (1-\theta)\hat{S}_0]}} \right) \right| \tag{A15}$$

$$= |\bar{L}(0)|. \tag{A16}$$

Equation (A14) follows by taking limits in (13), (A15) is a simple rearrangement, and (A16) follows from application of l'Hôpital's Rule. Standard results for the Lambert W function imply $\bar{L}(0) = 0$. *Q.E.D.*

Verification of Table 1 Entries

The equilibrium condition is $P_c^* = c$. Firms earn no profit under perfect competition: $\Pi_c^* = 0$. No consumers purchase in case (SR1), implying $Q_c^* = 0$. All susceptibles purchase in case (SR3), implying $Q_c^* = \hat{S}_0$. In case (SR2), Q_c^* can be found by substituting $P_c^* = c$ in equation (18).

To find $R_\infty(Q_c^*)$, note $R_\infty(Q_c^*) = 1 - I_\infty(Q_c^*) - S_\infty(Q_c^*) - \theta Q_c^* = 1 - S_\infty(Q_c^*) - \theta Q_c^*$ since $I_\infty(Q_c^*) = 0$. Substituting $Q_c^* = 0$ gives the entry for $R_\infty(Q_c^*)$ in case (SR1), and substituting $Q_c^* = \hat{S}_0$ gives the entry for $R_\infty(Q_c^*)$ in case (SR3). To find $R_\infty(Q_c^*)$ in case (SR2), set $c = MPB(Q_c^*)$ in equation (16) and rearrange, yielding

$$S_\infty(Q_c^*) = (1-\tilde{c})(\hat{S}_0 - \theta Q_c^*) \tag{A17}$$

Substituting (A17) into $R_\infty(Q_c^*) = 1 - S_\infty(Q_c^*) - \theta Q_c^*$ and rearranging yields $R_\infty(Q_c^*) = 1 - (1-\tilde{c})\hat{S}_0 - \tilde{c}\theta Q_c^*$. Substituting from the table entry for $Q_c^* = \hat{S}_0 \tilde{Q}_c^*$ yields the table entry for $R_\infty(Q_c^*)$.

Substituting $Q_c^* = 0$ in (16) gives MPB_c^* in case (SR1), and substituting $Q_c^* = \hat{S}_0$ in (16) gives MPB_c^* in case (SR3). For some but not all consumers to purchase in case (SR2) requires $MPB_c^* = c$.

Substituting $Q_c^* = 0$ in (23) gives MSB_c^* in case (SR1), and substituting $Q_c^* = \hat{S}_0$ in (23) gives MSB_c^* in case (SR3). Substituting from (A17) into (23) yields MSB_c^* in case (SR2).

The table entries for MEX_c^* can be obtained by subtracting other table entries: $MEX_c^* = MSB_c^* - MPB_c^*$. To derive the table table entries for W_c^*, by definition $W_c^* = SB_c^* - cQ_c^* = H[1-R_\infty(Q_c^*)] - cQ_c^*$, where the second equation follows from (21). Substituting other table entries into this equation gives the table entries for W_c^*. *Q.E.D.*

Proof of Proposition 1

Preliminaries: For conciseness, let $\mathcal{R}_0'$ denote the boundary value of $\mathcal{R}_0$ between cases (SR1) and (SR2) and $\mathcal{R}_0''$ between cases (SR2) and (SR3), i.e.,

$$\mathcal{R}_0' = \frac{|\ln(1-\tilde{c})|}{\hat{I}_0 + \tilde{c}\hat{S}_0} \tag{A18}$$

$$\mathcal{R}_0'' = \frac{|\ln(1-\tilde{c})|}{\hat{I}_0 + (1-\theta)\tilde{c}\hat{S}_0}. \tag{A19}$$

The subsequent analysis makes repeated use of the fact that

$$\frac{\partial S_\infty(Q)}{\partial \mathcal{R}_0} < 0. \tag{A20}$$

To see this, the Implicit Function Theorem can be applied to (12) to compute the derivative

$$\frac{\partial S_\infty(Q)}{\partial \mathcal{R}_0} = \frac{-S_\infty(Q)}{1-\mathcal{R}_0 S_\infty(Q)}[\hat{I}_0 + S_0(Q) - S_\infty(Q)]. \tag{A21}$$

The first factor is negative by Lemma 5. The factor in square brackets is positive since $\hat{I}_0 + S_0(Q) - S_\infty(Q) > S_0(Q) - S_\infty(Q) > 0$, where the first inequality follows from $\hat{I}_0 > 0$ and the second by Lemma 4.

Results for P_c^*, Π_c^*, and Q_c^*: The results for P_c^* and Π_c^* are obvious from Table 1. To show Q_c^* is weakly increasing, it can be verified that it is continuous at thresholds $\mathcal{R}_0'$ and $\mathcal{R}_0''$. In case (SR2), $\partial Q_c^*/\partial \mathcal{R}_0 = -\ln(1-\tilde{c})/\theta\tilde{c}\mathcal{R}_0^2 > 0$. Hence, Q_c^* is weakly increasing in $\mathcal{R}_0$ for all $\mathcal{R}_0 > 0$ and strictly increasing for $\mathcal{R}_0$ in the interior of case (SR2).

Results for MSB_c^*: To provide a roadmap for the analysis, we first look at cases (SR1) and (SR2) and show that MSB_c^* has a unique local maximum over $\mathcal{R}_0 \leq \mathcal{R}_0''$. Furthermore, this restricted local maximum is the restricted global maximum over $\mathcal{R}_0 \leq \mathcal{R}_0''$. We then look at cases (SR3) and (SR4) and show that MSB_c^* has at most one restricted local maximum over $\mathcal{R}_0 > \mathcal{R}_0''$. If no restricted local maximum exists there, then we show that the restricted maximum over $\mathcal{R}_0 \leq \mathcal{R}_0''$ is the global maximum over all $\mathcal{R}_0 > 0$. If a restricted local maximum exists in cases (SR3) and (SR4), then either it or the restricted local maximum over $\mathcal{R}_0 \leq \mathcal{R}_0''$ is the global maximum. This establishes, in sum, that MSB_c^* has at most two local maxima, one of which is the global maximum.

To prove that MSB_c^* has a unique local restricted maximum over $\mathcal{R}_0 \leq \mathcal{R}_0''$, we will show that MSB_c^* is increasing in a neighborhood around 0, quasiconcave for all $\mathcal{R}_0 \in (0, \mathcal{R}_0']$, continuous at $\mathcal{R}_0'$, and decreasing for all $\mathcal{R}_0 \in (\mathcal{R}_0', \mathcal{R}_0'')$. The arguments are made in reverse order. It is clear from

inspection of Table 1 that MSB_c^* is decreasing for $\mathcal{R}_0 \in (\mathcal{R}_0', \mathcal{R}_0'')$. To show MSB_c^* is continuous at $\mathcal{R}_0'$, using the table entry for MSB_c^* in case (SR1),

$$\lim_{\mathcal{R}_0 \uparrow \mathcal{R}_0'} MSB_c^* = \frac{\lim_{\mathcal{R}_0 \uparrow \mathcal{R}_0'} MPB_c^*}{1 - \mathcal{R}_0'(1 - \lim_{\mathcal{R}_0 \uparrow \mathcal{R}_0'} MPB_c^*/\theta H)\hat{S}_0} = \frac{\theta H \tilde{c}(\hat{I}_0 + \tilde{c}\hat{S}_0)}{\hat{I}_0 + \tilde{c}\hat{S}_0 + (1-\tilde{c})\hat{S}_0 \ln(1-\tilde{c})}. \tag{A22}$$

The first equality follows from substituting the table entry for MPB_c^* in case (SR1) directly as well as substituting the implication of that table entry that $S_\infty(0) = \hat{S}_0(1 - MPB_c^*/\theta H)$. The second equality follows from $\lim_{\mathcal{R}_0 \uparrow \mathcal{R}_0'} MPB_c^* = c$ by continuity and from substituting from (A18). Using the table entry for MSB_c^* in case (SR2),

$$\lim_{\mathcal{R}_0 \downarrow \mathcal{R}_0'} MSB_c^* = \frac{\theta H \tilde{c}^2}{\tilde{c} + (1-\tilde{c})[\ln(1-\tilde{c}) + \mathcal{R}_0'\hat{I}_0]} = \frac{\theta H \tilde{c}(\hat{I}_0 + \tilde{c}\hat{S}_0)}{\hat{I}_0 + \tilde{c}\hat{S}_0 + (1-\tilde{c})\hat{S}_0 \ln(1-\tilde{c})}. \tag{A23}$$

The equality of (A22) and (A23) proves the continuity of MSB_c^* at $\mathcal{R}_0'$.

We next show MSB_c^* is quasiconcave for all $\mathcal{R}_0$ in case (SR1). Differentiating the relevant table entry, substituting from (A21), and eliminating positive constants shows that $\partial MSB_c^*/\partial \mathcal{R}_0$ has the same sign as

$$\left[\hat{S}_0 - S_\infty(0)\right]\left[1 - \mathcal{R}_0 S_\infty(0)\right] + \left[\hat{I}_0 + \hat{S}_0 - S_\infty(0)\right]\left(1 - \mathcal{R}_0 \hat{S}_0\right). \tag{A24}$$

The second derivative of (A24) with respect to $\mathcal{R}_0$—after substituting from (A21), and rearranging considerably—can be shown to equal

$$2\frac{\partial S_\infty(0)}{\partial \mathcal{R}_0}\left[\hat{I}_0 + \hat{S}_0 - S_\infty(0)\right], \tag{A25}$$

which is negative—as can be shown using arguments similar to those behind (A20). Hence, (A24) is concave. In the limit $\mathcal{R}_0 \downarrow 0$, (A24) approaches $2[\hat{S}_0 - S_\infty(0)] + \hat{I}_0$, which is positive by Lemma 4 and $\hat{I}_0 > 0$. Having established that (A24) is concave throughout (SR1) and initially positive, we have that (A24) can change sign at most once. Therefore, $\partial MSB_c^*/\partial \mathcal{R}_0$ is either nonnegative throughout case (i) or positive then negative. In either event, this proves that MSB_c^* is quasiconcave in (SR1). We have already established MSB_c^* is increasing in a neighborhood of $\mathcal{R}_0$ above 0, the last step needed to prove that MSB_c^* has a unique restricted local maximum over $\mathcal{R}_0 \leq \mathcal{R}_0''$, which is a global maximum on that restricted set.

We next look at the behavior of MSB_c^* in cases (SR3) and (SR4), showing it has a most one restricted local maximum over $\mathcal{R}_0 > \mathcal{R}_0''$. Similar calculations used in the previous paragraph can be used here to establish the concavity of the following function,

$$\left[(1-\theta)\hat{S}_0 - S_\infty(\hat{S}_0)\right]\left[1 - \mathcal{R}_0 S_\infty(\hat{S}_0)\right] + \left[\hat{I}_0 + (1-\theta)\hat{S}_0 - S_\infty(\hat{S}_0)\right]\left[1 - (1-\theta)\mathcal{R}_0 \hat{S}_0\right], \tag{A26}$$

which determines the sign of $\partial MSB_c^*/\partial \mathcal{R}_0$ in (SR3) and (SR4). Thus, (A26) has at most two roots in those cases, which cannot both be local maxima, implying that MSB_c^* has at most one local maximum over $\mathcal{R}_0 > \mathcal{R}_0''$. The limit as $\mathcal{R}_0 \uparrow \infty$ of (A26) equals

$$\hat{I}_0 + 2(1-\theta)\hat{S}_0 - (1-\theta)\hat{S}_0[\hat{I}_0 + (1-\theta)\hat{S}_0] \lim_{\mathcal{R}_0 \uparrow \infty} \mathcal{R}_0 \tag{A27}$$

after substituting $\lim_{\mathcal{R}_0\uparrow\infty} S_\infty(\hat{S}_0) = \lim_{\mathcal{R}_0\uparrow\infty}[\mathcal{R}_0 S_\infty(\hat{S}_0)] = 0$ by Lemma 8. Expression (A27) approaches $-\infty$ since it involves $\mathcal{R}_0$ multiplied by negative constant. Since the limit $\mathcal{R}_0 \uparrow \infty$ cannot produce a restricted supremum over $\mathcal{R}_0 > \mathcal{R}_0''$, the restricted supremum is either the lower boundary of (SR3), i.e., $\mathcal{R}_0''$, or is the interior restricted maximum. If $\mathcal{R}_0''$ provides the restricted supremum over $\mathcal{R}_0 > \mathcal{R}_0''$, this cannot be a global maximum since MSB_c^* is decreasing in (SR2); the restricted maximum over $\mathcal{R}_0 \leq \mathcal{R}_0''$ must then be the global maximum.

Other Results: For space considerations, results for MPB_c^*, $R_\infty(Q_c^*)$, and MEX_c^* are relegated to Online Appendix B. The comparative statics for W_c^* are obvious from inspection of the table in view of (A21). *Q.E.D.*

Proof of Proposition 2

Results for MSB_m^*: We show that the limits of MSB_m^* for extreme values of $\mathcal{R}_0$ are exceed by interior values. We have

$$\lim_{\mathcal{R}_0\downarrow 0} MSB_m^* = \lim_{\mathcal{R}_0\downarrow 0} MSB_c^* = \theta H\left[1 - \frac{1}{\hat{S}_0}\lim_{\mathcal{R}_0\downarrow 0} S_\infty(0)\right] = \theta H\left(1 - \frac{\hat{S}_0}{\hat{S}_0}\right) = 0, \tag{A28}$$

where the first equality follows since $MSB_m^* = MSB_c^*$ for all $\mathcal{R}_0$ in case (SR1), the second equality follows from the entry for MSB_c^* in case (SR1) in Table 1, and the third equality follows from Lemma 7. To examine the upper limit, the proof of Proposition 1 showed that MSB_c^* asymptotes downward toward $\lim_{\mathcal{R}_0\uparrow\infty} MSB_c^* = 1$. Since $MSB_m^* = MSB_c^*$ in case (SR4), and all $\mathcal{R}_0$ above a sufficiently high value are contained in case (SR4), MSB_m^* must also slope downward toward its asymptote. Thus MSB_m^* is higher at interior values of $\mathcal{R}_0$ than the extremes.

Other Results: For space considerations, results for Π_m^*, $R_\infty(Q_m^*)$, and MEX_m^* are relegated to Online Appendix B. *Q.E.D.*

Proof of Proposition 3

The sketch of the proof in the text omitted two details filled in here. We first prove Q^{**} for $\mathcal{R}_0$ in a neighborhood above 0. Taking limits in (23),

$$\lim_{\mathcal{R}_0\downarrow 0} MSB(Q) = \theta H\left[1 - \frac{S_0(Q)}{S_0(Q)}\right] = 0. \tag{A29}$$

Hence, there exists $\mathcal{R}_0$ in a neighborhood above 0 and $\epsilon \in (0, c)$ such that $MSB(Q) < \epsilon$. For $\mathcal{R}_0$ in this neighborhood, $W(Q) = \int_0^Q [MSB(x) - c]dx < (\epsilon - c)Q < 0 = W(0)$. Thus, $Q^{**} = 0$ for $\mathcal{R}_0$ in this neighborhood.

We next prove $GS^{**} > 0$ for some $\mathcal{R}_0 \in (0, \infty)$. Since $Q^{**} = 0$ for all $\mathcal{R}_0$ in neighborhood of 0, $Q^* \leq Q^{**} = 0$ implies $Q^* = 0$ for all $\mathcal{R}_0$ in a neighborhood of 0. The text argued $Q_m^* = \hat{S}_0$ for sufficiently high $\mathcal{R}_0$, implying $\hat{S}_0 = Q_m^* \leq Q^{**} \leq hS_0$, implying $Q^{**} = \hat{S}_0$ for sufficiently high $\mathcal{R}_0$. By the Theorem of the Maximum, since Q^{**} is a maximizer of continuous function $W(Q)$, Q^{**} is continuous, implying the existence of $\mathcal{R}_0 \in (0, \infty)$ such that $Q^* \in (0, \hat{S}_0)$. This Q^{**} must satisfy the

first-order condition $MSB(Q^{**}) = c$, implying $MPB(Q^{**}) + MEX(Q^{**}) = c$, implying $MPB(Q^{**}) < c$ since $MEX(Q) > 0$ for all $Q \in (0, \hat{S}_0)$ by (24). *Q.E.D.*

Proof of Proposition 6

Suppose $\mathcal{R}_0 \hat{S}_0 > 2$. Then $1 < \mathcal{R}_0 \hat{S}_0/2 < \mathcal{R}_0[\hat{S}_0 + S_\infty(0)]/2 = \mathcal{R}_0[S_0(0) + S_\infty(0)]/2$, where the second step follows from $S_\infty(0) > 0$ by Lemma 4. This chain of inequalities implies that (33) holds at $Q = 0$ and thus that the vaccine exhibits initially increasing social returns.

At a general output level $Q \in (0, \hat{S}_0)$,

$$\mathcal{R}_0 \left[\frac{S_0(Q) + S_\infty(Q)}{2} \right] > \mathcal{R}_0 \left(\frac{S_0(Q)}{2} \right) = \mathcal{R}_0 \left(\frac{\hat{S}_0 - \theta Q}{2} \right) \geq \mathcal{R}_0 \left(\frac{(1-\theta)\hat{S}_0}{2} \right). \tag{A30}$$

If $\mathcal{R}_0 \hat{S}_0 \geq 2/(1-\theta)$, then the last expression weakly exceeds 1, implying (33) holds for all feasible Q, implying the vaccine exhibits everywhere increasing social returns. *Q.E.D.*

Proof of Proposition 7

Universal vaccination with a perfectly effective vaccine implies $S_0(\hat{S}_0) = (1-\theta)\hat{S}_0 = 0$. We thus have

$$\mathcal{R}_0 \left[\frac{S_0(\hat{S}_0) + S_\infty(\hat{S}_0)}{2} \right] = \mathcal{R}_0 \left(\frac{S_\infty(\hat{S}_0)}{2} \right) < \frac{1}{2}. \tag{A31}$$

Hence, (33) does not hold for $Q = \hat{S}_0$, implying the vaccine does not exhibit increasing social returns for $Q = \hat{S}_0$ by Proposition 5. *Q.E.D.*

Proof of Proposition 8

The assumption $c = 0$ implies $\tilde{c} = 0$, leaving two cases in Table B1: (SR2)–(SR3) and (SR4). Nesting those cases, we can write

$$\Delta\Pi_m^* = \theta H \left\{ \hat{I}_0 + \hat{S}_0 \Phi_I(0) - Q_{mv}^* \Phi_I(Q_{mv}^*) \right\}, \tag{A32}$$

where Q_{mv}^* solves $\max_{Q \in [0, \hat{S}_0]} Q\Phi_I(Q)$. Since $Q_{mv}^* > 0$, we have $Q_{mv}^* \Phi_I(Q_{mv}^*) < Q_{mv}^* \Phi_I(0) \leq \hat{S}_0 \Phi_I(0)$, where the first inquality follows from $\partial \Phi_I(Q)/\partial Q < 0$ by (B28) and the second inequality from $Q_{mv}^* \in [0, \hat{S}_0]$. Substituting the preceding inequality into (A32) yields $\Delta\Pi_m^* > \theta H \hat{I}_0$. Thus, $\Delta\Pi_m^* > 0$ for all $\mathcal{R}_0 > 0$.

To derive the results on limits of $\Delta\Pi_m^*$, we have that $\lim_{\mathcal{R}_0 \downarrow 0} \Phi_I(Q) = \lim_{\mathcal{R}_0 \downarrow 0}[1 - S_\infty(Q)/S_0(Q)] = 1$ for all $Q \in [0, \hat{S}_0]$ since $\lim_{\mathcal{R}_0 \downarrow 0} S_\infty(Q) = \hat{S}_0 - \theta Q = S_0(Q)$ by Lemma 7. Hence, $\lim_{\mathcal{R}_0 \downarrow 0} \Delta\Pi_m^* = \theta H \hat{I}_0$. For all $Q \in [0, \hat{S}_0]$, $\lim_{\mathcal{R}_0 \uparrow \infty} \Phi_I(Q) = 1$ since $\lim_{\mathcal{R}_0 \uparrow \infty} S_\infty(Q) = 0$ by Lemma 8. Therefore,

$$\lim_{\mathcal{R}_0 \uparrow \infty} Q_{mv}^* \Phi_I(Q_{mv}^*) = \lim_{\mathcal{R}_0 \uparrow \infty} \left\{ \max_{Q \in [0, \hat{S}_0]} Q\Phi_I(Q) \right\} = \max_{Q \in [0, \hat{S}_0]} \left[Q \lim_{\mathcal{R}_0 \uparrow \infty} \Phi_I(Q) \right] = \hat{S}_0 \cdot 1. \tag{A33}$$

Substituting from (A33) into (A32) along with $\lim_{\mathcal{R}_0\uparrow\infty}\Phi_I(0)=1$ yields $\lim_{\mathcal{R}_0\uparrow\infty}\Delta\Pi_m^*=\theta H\hat{I}_0$. Now $\Delta\Pi_m^*>\theta H\hat{I}_0$ for all $\mathcal{R}_0>0$ implies $\theta H\hat{I}_0\leq\inf_{\mathcal{R}_0>0}\Delta\Pi_m^*\leq\lim_{\mathcal{R}_0\downarrow 0}\Delta\Pi_m^*=\theta H\hat{I}_0$, which in turn implies $\inf_{\mathcal{R}_0>0}\Delta\Pi_m^*=\theta H\hat{I}_0$.

Combining the results from the previous paragraph, $\lim_{\mathcal{R}_0\downarrow 0}\Delta\Pi_m^*=\lim_{\mathcal{R}_0\uparrow\infty}\Delta\Pi_m^*=\inf_{\mathcal{R}_0>0}=\theta H\hat{I}_0$. But the first paragraph showed $\Delta\Pi_m^*>\theta H\hat{I}_0$. Hence, $\Delta\Pi_m^*$ must be nonmonotonic in $\mathcal{R}_0$, higher in the interior than for either limiting value of $\mathcal{R}_0$.

Turning to limiting values of ΔW_m^* as $\mathcal{R}_0\downarrow 0$ and $\mathcal{R}_0\uparrow\infty$, one can show that (28) holds in these limits. Thus, the relevant case for computing W_{mv}^* is (SR4). Substituting $\tilde{c}=0$ into the relevant entry of Table 2 and multiplying by $\theta H\hat{S}_0$ to reverse the rescaling yields $W_{mv}^*=H[S_\infty(\hat{S}_0)+\theta\hat{S}_0]$. Subtracting from (35) and rearranging yields

$$\Delta W_m^*=H\left[\theta\hat{I}_0+(1-\theta)S_\infty(0)-S_\infty(\hat{S}_0)\right]. \tag{A34}$$

By Lemma 7, $\lim_{\mathcal{R}_0\downarrow 0}[(1-\theta)S_\infty(0)]=(1-\theta)\hat{S}_0$. The lemma also implies $\lim_{\mathcal{R}_0\downarrow 0}S_\infty(\hat{S}_0)=(1-\theta)\hat{S}_0$. Substituting these limits into (A34) yields $\lim_{\mathcal{R}_0\downarrow 0}\Delta W_m^*=\theta H\hat{I}_0$. By Lemma 8, $\lim_{\mathcal{R}_0\uparrow\infty}S_\infty(0)=\lim_{\mathcal{R}_0\uparrow\infty}S_\infty(\hat{S}_0)=0$. Substituting these limits into (A34) yields $\lim_{\mathcal{R}_0\uparrow 0}\Delta W_m^*=\theta H\hat{I}_0$.

The final step is to provide parameters for which $\Delta W_m^*<0$. Using Matlab, we verified that for $\mathcal{R}_0=2$, $\theta=0.5$, $\hat{I}_0=0.1$, $\hat{S}_0=0.8$, (28) holds, implying that the vaccine monopolist supplies first-best quantity $\hat{S}_0$, putting us in case (SR4). Subtracting the relevant Table 2 entry from (35) and simplifying yields $\Delta W_m^*=H[(1-\theta)S_\infty(0)+\theta\hat{I}_0-S_\infty(\hat{S}_0)]$, which Matlab calculations show equals –0.09 for the specified parameters. *Q.E.D.*

References

Adida, Elodie, Debabrata Dey, and Hamed Mamani. (2013) "Operational Issues and Network Effects in Vaccine Markets," *Europoean Journal of Operational Research* 231: 414–427.

Aguirregabiria, Victor, Jiaying Gu, Yao Luo, and Pedro Mira (2020) "A Dynamic Structural Model of Virus Diffusion and Network Production: A First Report, CEPR discussion paper no. DP14750.

Ahuja, Amrita, Susan Athey, Arthur Baker, Eric Budish, Juan C. Castillo, Rachel Glennerster, Scott D. Kominers, Michael Kremer, Jean N. Lee, Canice Prendergast, Christopher M. Snyder, Alex Tabarrok, Brandon J. Tan, and Witold Wiecek. (2021) "Preparing for a Pandemic: Accelerating Vaccine Availability," *American Economic Association Papers and Proceedings* 111: 331–335.

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

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

Atkeson, Andrew, Karen Kopecky, and Tao Zha. (2020) "Estimating and Forecasting Disease Scenarios for COVID-19 with an SIR Model," National Bureau of Economic Research working paper no. 27335.

Bauch, Chris T., and David J. D. Earn. (2004) "Vaccination and the Theory of Games," *Proceedings of the National Academy of Sciences* 101(36): 13391-13394

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

Bisin, Alberto and Andrea Moro. (2020) "Learning Epidemiology by Doing: The Empirical Implications of a Spatial SIR Model with Behavioral Responses," SSRN working paper no. 3625361.

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

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

Buckner, Jack H., Gerardo Chowell, and Michael R. Springborn. (2021) "Dynamic Prioritization of COVID-19 Vaccines When Social Distancing is Limited for Essential Workers," *Proceedings of the National Academy of Sciences* 118: .

Castillo, Juan C., Amrita Ahuja, Susan Athey, Arthur Baker, Eric Budish, Tasneem Chipty, Rachel Glennerster, Scott D. Kominers, Michael Kremer, Greg Larson, Jean Lee, Canice Prendergast, Christopher M. Snyder, Alex Tabarrok, Brandon J. Tan, and Witold Wiecek. (2021) "Market Design to Accelerate COVID-19 Vaccine Supply," *Science* 371(6534): 1107–1109.

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

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

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

Diekmann, O., J. A. P. Heesterbeek, and J. A. J. Metz. (1990) "On the Definition and the Computation of the Basic Reproduction Ratio R_0 in Models for Infections Diseases in Heterogeneous Populations," *Journal of Mathematical Biology* 28: 365–382.

Eichenbaum, Martin S., Sergio Rebelo, and Mathias Trabandt. (2020) "The Macroeconomics of Epidemics," National Bureau of Economics working paper no. 26882.

Enayati, Shakiba and Osman Y. Özaltin. (2020) "Optimal Influenza Vaccine Distribution with Equity," *European Journal of Operational Research* 283: 714–725.

Farboodi, Maryam, Gregor Jarosch, and Robert Shimer. (2020) " Internal and External Effects of Social Distancing in a Pandemic," National Bureau of Economic Research working paper no. 27059.

Fenichel, Eli P. (2013) "Economic Considerations for Social Distancing and Behavioral Based Policies During an Epidemic," *Journal of Health Economics* 32: 440–451.

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

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

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

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

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

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

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

Greenwood, Jeremy, Philipp Kircher, Cezar Santos, and Michèle Tertilt. (2019) "An Equilibrium Model of the African HIV/AIDS Epidemic," *Econometrica* 87: 1081–1113.

Hanlon Peter, Fergus Chadwick, Anoop Shah, Rachel Wood, Jon Minton, Gerry McCartney, Colin Fischbacher, Frances Mair, Dirk Jusmeier, Jason Matthiopoulos, and David A. McAllister. (2021) "COVID-19: Exploring the Implications of Long-term Condition Type and Extent of Multimorbitity on Years of Life Lost: A Modelling Study," *Wellcome Open Research* 5:75.

Jones, Callum J., Thomas Philippon, and Venky Venkateswaran. (2020) "Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home," National Bureau of Economic Research working paper no. 26984.

Keeling, Matt J. and J. V. Ross. (2015) "Optimal Prophylactic Vaccination in Segregated Populations: When Can We Improve on the Equalising Strategy?" *Epidemics* 11: 7–13.

Keeling, Matt J. and Andrew Shattock. (2012) "Optimal but Inequitable Prophylactic Distribution of Vaccine," *Epidemic* 4: 78–85.

Kermack, William O. and A. G. McKendrick. (1927) "A Contribution to the Mathematical Theory of Epidemics," *Proceedings of the Royal Society A (Mathematical, Physical and Engineering Sciences)* 115: 700–721.

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

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

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

Mamani, Hamed, Elodie Adida, and Debabrata Dey. (2012) "Vaccine Market Coordination Using Subsidy," *IIE Transactions on Healthcare Systems Engineering* 2: 78–96.

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

Manski, Charles F. (2010) "Vaccination with Partial Knowledge of External Effectiveness," *Proceedings of the National Academy of Sciences* 107: 3953–3960.

Manski, Charles F. (2017) "Mandating Vaccination with Unknown Indirect Effects," *Journal of Public Economic Theory* 19: 603–619.

Manski, Charles F. (2021) "Vaccination Planning Under Uncertainty, with Application to Covid-19," National Bureau of Economic Research working paper no. 28446.

Marseille Elliott, Bruce Larson, Dhruv S. Kazi, James G. Kahn, and Sydney Rosen. (2015) "Thresholds for the Cost-effectiveness of Interventions: Alternative Approaches," *Bulletin of the World Health Organization* 93(2): 118—124.

Martcheva, Maia. (2015) *An Introduction to Mathematical Epidemiology*. New York: Springer.

May, Robert. (2000) "Simple Rules with Complex Dynamics," *Science* 287(5453): 601–602.

McAdams, David. (2020) "Nash SIR: An Economic-Epidemiological Model of Strategic Behavior During a Viral Epidemic," SSRN working paper no. 393272.

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

Nguyen, Chantal and Jean M. Carlson. (2016) 'Optimizing Real-Time Vaccine Allocation in a Stochastic SIR Model," *PLoS One* 11(4): e0152950.

Public Health England. (2021) "COVID-19 Vaccine Surveillance Report (Week 29)." [Internet.] Downloaded July 29, 2021 from https://www.gov.uk/government/publications/covid-19-vaccine-surveillance-report.

Rachel, Lukasz. (2020) "An Analytical Model of Covid-19 Lockdown," London School of Economics Center for Macroeconomics discussion paper no. 2020–29.

Snyder, Christopher M., Kendall Hoyt, Dimitrios Gouglas, Thomas Johnston, and James Robinson. (2020) "Designing Pull Funding for a COVID-19 Vaccine," *Health Affairs* 39(9): 1633–1642.

Toxvaerd, Flavio. (2019) "Rational Disinhibition and Externalities in Prevention," *International Economic Review* 60: 1737–1755.

Toxvaerd, Flavio. (2020) "Equilibrium Social Distancing," University of Cambridge working paper no. 2020/08.

U.K. Government Office for Science. (2020) "The R Number and Growth Rate in the UK." [Internet.] Downloaded September 28, 2020 from https://www.gov.uk/guidance/the-r-number-in-the-uk#latest-r-number-and-growth-rate.

U.K. Office for National Statistics. (2020) "Coronavirus (COVID-19) Infection Survey Pilot: England, Wales and Northern Ireland, 25 September 2020." [Internet.] Downloaded September 28, 2020 from https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/englandwalesandnorthernireland25september2020#number-of-people-in-england-who-had-covid-19.

Ward, Courtney J. (2014) "Influenza Vaccination Campaigns: Is an Ounce of Prevention Worth a Pound of Cure?" *American Economic Journal: Applied Economics* 6: 38–72.

Table 1: Equilibrium Variables under Perfect Competition as Functions of $\mathcal{R}_0$

	Case		
	(SR1)	(SR2)	(SR3), (SR4)
Variable	$\mathcal{R}_0 \in \left(0, \frac{\lvert\ln(1-\tilde{c})\rvert}{\hat{I}_0+\tilde{c}\hat{S}_0}\right]$	$\mathcal{R}_0 \in \left(\frac{\lvert\ln(1-\tilde{c})\rvert}{\hat{I}_0+\tilde{c}\hat{S}_0}, \frac{\lvert\ln(1-\tilde{c})\rvert}{\hat{I}_0+(1-\theta)\tilde{c}\hat{S}_0}\right]$	$\mathcal{R}_0 \in \left(\frac{\lvert\ln(1-\tilde{c})\rvert}{\hat{I}_0+(1-\theta)\tilde{c}\hat{S}_0}, \infty\right)$
P_c^*	c	c	c
Q_c^*	0	$\frac{1}{\theta}\left\{\hat{S}_0+\frac{1}{\tilde{c}}\left[\frac{1}{\mathcal{R}_0}\ln(1-\tilde{c})+\hat{I}_0\right]\right\}$	$\hat{S}_0$
Π_c^*	0	0	0
$R_\infty(Q_c^*)$	$1-S_\infty(0)$	$1-\hat{S}_0-\hat{I}_0+\frac{1}{\mathcal{R}_0}\lvert\ln(1-\tilde{c})\rvert$	$1-S_\infty(\hat{S}_0)-\theta\hat{S}_0$
MPB_c^*	$\theta H\Phi_I(0)$	c	$\theta H\Phi_I(\hat{S}_0)$
MSB_c^*	$\frac{\theta H\Phi_I(0)}{1-\mathcal{R}_0 S_\infty(0)}$	$\frac{\theta H\tilde{c}^2}{\tilde{c}+(1-\tilde{c})[\ln(1-\tilde{c})+\mathcal{R}_0\hat{I}_0]}$	$\frac{\theta H\Phi_I(\hat{S}_0)}{1-\mathcal{R}_0 S_\infty(\hat{S}_0)}$
MEX_c^*	$\frac{\theta H\Phi_I(0)\mathcal{R}_0 S_\infty(0)}{1-\mathcal{R}_0 S_\infty(0)}$	$\frac{\theta H\tilde{c}(1-\tilde{c})[\lvert\ln(1-\tilde{c})\rvert-\mathcal{R}_0\hat{I}_0]}{\tilde{c}+(1-\tilde{c})[\ln(1-\tilde{c})+\mathcal{R}_0\hat{I}_0]}$	$\frac{\theta H\Phi_I(\hat{S}_0)\mathcal{R}_0 S_\infty(\hat{S}_0)}{1-\mathcal{R}_0 S_\infty(\hat{S}_0)}$
W_c^*	$HS_\infty(0)$	$H\hat{S}_0(1-\tilde{c})$	$H\left[\theta\hat{S}_0(1-\tilde{c})+S_\infty(\hat{S}_0)\right]$

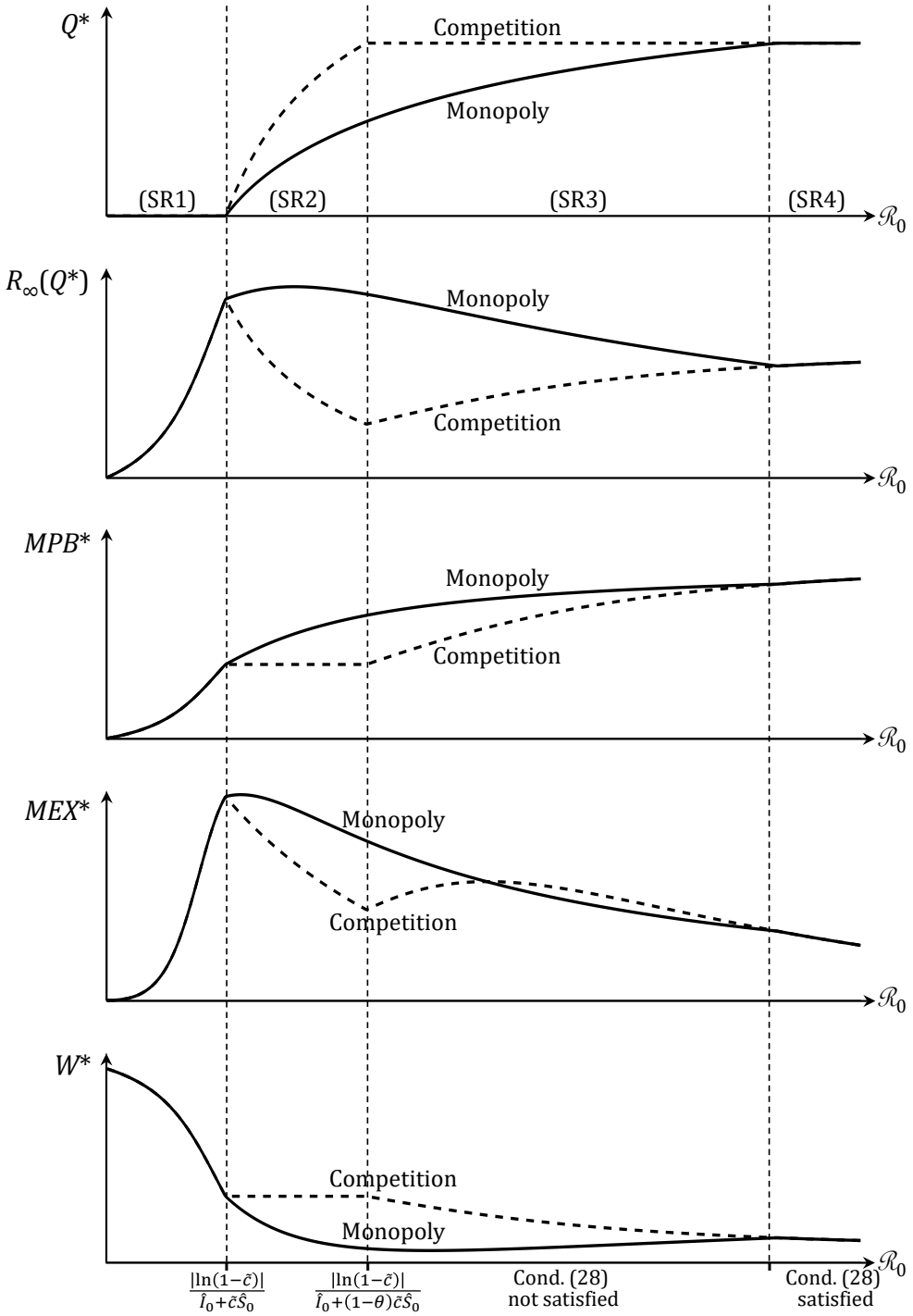
Notes: Computable expressions for $S_\infty(0)$ and $S_\infty(\hat{S}_0)$ can be derived from equation (13): defining $\hat{\epsilon}_I = e^{-\mathcal{R}_0\hat{I}_0}$ and $\hat{\epsilon}_S = e^{-\mathcal{R}_0\hat{S}_0}$, we have $S_\infty(0) = \lvert\bar{L}(-\mathcal{R}_0\hat{S}_0\hat{\epsilon}_I\hat{\epsilon}_S)\rvert/\mathcal{R}_0$ and $S_\infty(\hat{S}_0) = \lvert\bar{L}(-\mathcal{R}_0(1-\theta)\hat{S}_0\hat{\epsilon}_I\hat{\epsilon}_S^{1-\theta})\rvert/\mathcal{R}_0$. Infection probabilities can be computed using (15): $\Phi_I(0) = 1-S_\infty(0)/\hat{S}_0$ and $\Phi_I(\hat{S}_0) = 1-S_\infty(\hat{S}_0)/(1-\theta)\hat{S}_0$. The distinction between cases (SR3) and (SR4) in the last column, relevant for monopoly in the next table, is irrelevant for perfect competition here.

Table 2: Equilibrium Variables under Monopoly as Functions of $\mathcal{R}_0$

	Case		
	(SR1)	(SR2), (SR3)	(SR4)
Variable	$\mathcal{R}_0 \in \left(0, \frac{\lvert\ln(1-\tilde{c})\rvert}{\hat{I}_0+\tilde{c}\hat{S}_0}\right]$	$\mathcal{R}_0 > \lvert\ln(1-\tilde{c})\rvert/(\hat{I}_0+\tilde{c}\hat{S}_0)$ but does not satisfy (28)	$\mathcal{R}_0$ satisfies (28)
P_m^*	†	$\theta H\left[1-\frac{S_\infty(Q_m^*)}{\hat{S}_0-\theta Q_m^*}\right]$	$\theta H\left[1-\frac{S_\infty(\hat{S}_0)}{(1-\theta)\hat{S}_0}\right]$
Q_m^*	0	‡	$\hat{S}_0$
Π_m^*	0	$\theta H\left[1-\frac{S_\infty(Q_m^*)}{\hat{S}_0-\theta Q_m^*}-\tilde{c}\right]Q_m^*$	$\theta H\left[(1-\tilde{c})\hat{S}_0-\frac{S_\infty(\hat{S}_0)}{(1-\theta)}\right]$
$R_\infty(Q_m^*)$	$1-S_\infty(0)$	$1-S_\infty(Q_m^*)-\theta Q_m^*$	$1-S_\infty(\hat{S}_0)-\theta\hat{S}_0$
MPB_m^*	$\theta H\Phi_I(0)$	$\theta H\Phi_I(Q_m^*)$	$\theta H\Phi_I(\hat{S}_0)$
MSB_m^*	$\frac{\theta H\Phi_I(0)}{1-\mathcal{R}_0 S_\infty(0)}$	$\frac{\theta H\Phi_I(Q_m^*)}{1-\mathcal{R}_0 S_\infty(Q_m^*)}$	$\frac{\theta H\Phi_I(\hat{S}_0)}{1-\mathcal{R}_0 S_\infty(\hat{S}_0)}$
MEX_m^*	$\frac{\theta H\Phi_I(0)\mathcal{R}_0 S_\infty(0)}{1-\mathcal{R}_0 S_\infty(0)}$	$\frac{\theta H\Phi_I(Q_m^*)\mathcal{R}_0 S_\infty(Q_m^*)}{1-\mathcal{R}_0 S_\infty(Q_m^*)}$	$\frac{\theta H\Phi_I(\hat{S}_0)\mathcal{R}_0 S_\infty(\hat{S}_0)}{1-\mathcal{R}_0 S_\infty(\hat{S}_0)}$
W_m^*	$HS_\infty(0)$	$H\left[(1-\tilde{c})\theta\tilde{Q}_m^*+S_\infty(Q_m^*)\right]$	$H\left[\theta\hat{S}_0(1-\tilde{c})+S_\infty(\hat{S}_0)\right]$

Notes: Equations (13) and (15) provide formulas for computing $S_\infty(Q_m^*)$ and $\Phi_I(Q_m^*)$, respectively, given Q_m^*. See the previous table for simpler expressions for $S_\infty(0)$, $S_\infty(\hat{S}_0)$, $\Phi_I(0)$, and $\Phi_I(\hat{S}_0)$. The distinction between cases (SR2) and (SR3) in the middle column, relevant for perfect competition in previous table, is irrelevant for monopoly here. The set of $\mathcal{R}_0$ satisfying equation (28) need not form an interval. †Any value $P_m^* \geq c$ is consistent with zero sales in equilibrium. ‡Entry is the Q_m^* solving (27).

Figure 1: Graphs of Selected Equilibrium Variables as Functions of $\mathcal{R}_0$



Notes: Graph of formulas provided in Tables 1 and 2, illustrated for specific parameter values ($\theta = 0.7$, $\alpha = \mu = 0$, $c = 0.3$, $H = 1$, $\hat{I}_0 = 0.1$, $\hat{S}_0 = 0.8$). Where the dotted and solid curves overlap, the solid curve represents both industry structures. Panel illustrating equilibrium prices omitted since $P_m^* = MPB_m^*$ in all cases in which the monopolist serves some consumers and since $P_c^* = c$ does not require illustration. Also omitted is a panel illustrating equilibrium profit since the graph for Π_m^* resembles that for Q_m^* or MPB_m^* and since $\Pi_c^* = 0$ does not require illustration.

Optimal Vaccine Subsidies for Epidemic Diseases
Online Appendixes

Matthew Goodkin-Gold
Michael Kremer
Christopher M. Snyder
Heidi Williams

This document contains four online appendixes, not for publication, providing further analytical details or extensions omitted from the published paper for space considerations. The proofs in Appendix A accompanying the article are fairly complete but still omit certain cases or technical details for space considerations. That material is provided in Online Appendix B1. Online Appendix B2 analyzes Cournot competition among n firms. This analysis nests perfect competition studied in the article in the limit $n \uparrow \infty$ and also nests monopoly studied in the article setting $n = 1$. Online Appendix B3 extends the analysis of homogeneous consumers to allow consumers to vary in disease harm H_i. Online Appendix B4 repeats all of the analysis that was undertaken in the article in a short-run model with a fixed population and one-shot vaccination campaign, now for a long-run model with population flows reflecting births and deaths, examining the continuous rate of vaccination in the steady-state equilibrium of that model.

Appendix B1. Further Proof Details

Completing Proof of Proposition 1

It remains to provide results for MPB_c^*, $R_\infty(Q_c^*)$, and MEX_c^*, omitted from the proof in Appendix A.

Results for MPB_c^*: To show MPB_c^* is weakly increasing, start with case (SR1). Differentiating the table entry,

$$\frac{\partial MPB_c^*}{\partial \mathcal{R}_0} = -\left(\frac{\theta H}{\hat{S}_0}\right)\frac{\partial S_\infty(0)}{\partial \mathcal{R}_0}. \tag{B1}$$

By (A20), $\partial S_\infty(Q)/\partial \mathcal{R}_0 < 0$ for all $Q \in [0, \hat{S}_0]$, including $Q_c^* = 0$, implying (B1) is positive. In case (SR2), MPB_c^* is constant. Differentiating the table entry in cases (SR3) and (SR4),

$$\frac{\partial MPB_c^*}{\partial \mathcal{R}_0} = -\left[\frac{\theta H}{(1-\theta)\hat{S}_0}\right]\frac{\partial S_\infty(\hat{S}_0)}{\partial \mathcal{R}_0}, \tag{B2}$$

which is negative since $\partial S_\infty(Q)/\partial \mathcal{R}_0 < 0$ by (A20) for all $Q \in [0, \hat{S}_0]$, including $Q_c^* = \hat{S}_0$.

The last step in deriving comparative statics for MPB_c^* is to show it is continuous at both endpoints of case (SR2). Now $MPB(Q)$ is continuous in Q because it is differentiable in Q by (17). Further, $MPB(Q)$ is continuous in $\mathcal{R}_0$ because $S_\infty(Q)$ is differentiable in $\mathcal{R}_0$ by (A21). Since Q_c^* is continuous at both endpoints of case (SR2) as argued in the first paragraph of this proof, we have that MPB_c^* is continous at $\mathcal{R}_0'$ and $\mathcal{R}_0''$.

Results for $R_\infty(Q_c^*)$: To derive the comparative statics for $R_\infty(Q_c^*)$, combining the table entries with (A20) shows $R_\infty(Q_c^*)$ is increasing in $\mathcal{R}_0$ in case (SR1) as well as cases (SR3) and (SR4). The table entry is obviously decreasing in $\mathcal{R}_0$ in case (SR2). We thus have that $R_\infty(Q_c^*)$ attains a local maximum at $\mathcal{R}_0'$ if we can establish that $R_\infty(Q_c^*)$ is continuous at $\mathcal{R}_0'$. Using the table entry for $R_\infty(Q_c^*)$ in case (SR1),

$$\lim_{\mathcal{R}_0\uparrow\mathcal{R}_0'} R_\infty(Q_c^*) = 1 - \lim_{\mathcal{R}_0\uparrow\mathcal{R}_0'} S_\infty(0) = 1 - \left(1 - \lim_{\mathcal{R}_0\uparrow\mathcal{R}_0'} \frac{MPB_c^*}{\theta H}\right)\hat{S}_0 = 1-(1-\tilde{c})\hat{S}_0. \tag{B3}$$

The second equality follows from the table entry for MPB_c^* in case (SR1): $MPB_c^* = \theta H \Phi_I(0) = \theta H[1 - S_\infty(0)/\hat{S}_0]$ by (15). The third equality follows from the continuity of MPB_c^* at $\mathcal{R}_0'$, allowing us to substitute the table entry for $MPB_c^* = c$ in case (SR2). Using the table entry for $R_\infty(Q_c^*)$ in case (SR2),

$$\lim_{\mathcal{R}_0\downarrow\mathcal{R}_0'} R_\infty(Q_c^*) = 1 - \hat{S}_0 - \hat{I}_0 + \frac{1}{\mathcal{R}_0'}|\ln(1-\tilde{c})| = 1-(1-\tilde{c})\hat{S}_0. \tag{B4}$$

The equality between (B3) and (B4) proves the continuity of $R_\infty(Q_c^*)$ at $\mathcal{R}_0'$.

Since $R_\infty(Q_c^*)$ is increasing in $\mathcal{R}_0$ in cases (SR3) and (SR4), the other candidate for a supremum is

$$\lim_{\mathcal{R}_0\uparrow\infty} R_\infty(Q_c^*) = 1 - \theta\hat{S}_0. \tag{B5}$$

This equality follows from taking the limit $\mathcal{R}_0 \uparrow \infty$ of the table entry in cases (SR3) and (SR4) and noting that $\lim_{\mathcal{R}_0\uparrow\infty} S_\infty(\hat{S}_0) = 0$ by Lemma 8. The local maximum is thus a global maximum if and only if $1-(1-\tilde{c})\hat{S}_0 \geq 1-\theta\hat{S}_0$. Rearranging gives $\tilde{c} \geq 1-\theta$.

Results for MEX_c^*: We use the same roadmap for the comparative-statics analysis of MEX_c^* as for MSB_c^*. We begin by proving that MEX_c^* has a unique local restricted maximum over $\mathcal{R}_0 \leq \mathcal{R}_0''$. We do this by showing that MEX_c^* is increasing in a neighborhood around 0, quasiconcave for all $\mathcal{R}_0 \in (0, \mathcal{R}_0']$, continuous at $\mathcal{R}_0'$, and decreasing for all $\mathcal{R}_0 \in (\mathcal{R}_0', \mathcal{R}_0'')$. The arguments are made in reverse order. Differentiating the table entry for case (SR2) yields

$$\frac{\partial MEX_c^*}{\partial \mathcal{R}_0} = \frac{-\theta H \tilde{c}^2(1-\tilde{c})\hat{I}_0}{\left\{\tilde{c} + (1-\tilde{c})[\ln(1-\tilde{c}) + \mathcal{R}_0\hat{I}_0]\right\}^2}, \tag{B6}$$

which is negative. The proof that MEX_c^* is continuous at $\mathcal{R}_0'$ is similar to that for MSB_c^* and omitted.

We next show MEX_c^* is quasiconcave for all $\mathcal{R}_0$ in case (SR1). Differentiating the relevant table entry, substituting from equation (A21), and eliminating positive constants shows that $\partial MEX_c^*/\partial \mathcal{R}_0$ has the same sign as

$$\left[\hat{S}_0 - S_\infty(0)\right]\left[1 - \mathcal{R}_0(\hat{I}_0 + \hat{S}_0)\right] + \mathcal{R}_0 S_\infty(0)\left[1 - \mathcal{R}_0 S_\infty(0)\right]\left[\hat{I}_0 + \hat{S}_0 - S_\infty(0)\right]. \tag{B7}$$

All of the factors in (B7) are definitively positive except for $1-\mathcal{R}_0(\hat{I}_0+\hat{S}_0)$. If this is also nonnegative, then $\partial MEX_c^*/\partial \mathcal{R}_0$ is positive in (SR1), implying MEX_c^* is quasiconcave in (SR1), as desired.

So suppose instead that

$$\mathcal{R}_0(\hat{I}_0 + \hat{S}_0) > 1. \tag{B8}$$

We will show that (B8) implies that (B7) is concave. The second derivative of (B7) with respect to $\mathcal{R}_0$—after substituting from (A21), rearranging considerably, and removing positive factors—can be shown to have the same sign as

$$S_\infty(0)\left[1-\mathcal{R}_0(\hat{I}_0+\hat{S}_0)\right]-\left[1-\mathcal{R}_0 S_\infty(0)\right](\hat{I}_0+\hat{S}_0)-S_\infty(0)\left[2(\hat{I}_0+\hat{S}_0)-S_\infty(0)\right] \\ -S_\infty(0)\left\{1-\mathcal{R}_0\left[\hat{I}_0+\hat{S}_0-S_\infty(0)\right]\right\}\left[\hat{I}_0+\hat{S}_0-2S_\infty(0)\right]. \quad \text{(B9)}$$

By (B8) and familiar arguments, all the terms in (B9) are negative except possibly the last. If the last term is also nonpositive, the whole expression is negative, establishing (B7) is concave. So suppose instead that the last term is positive. For this to be the case, one of its last two factors must be positive and the other negative. That is, one of the following two sets of conditions must hold:

$$1-\mathcal{R}_0\left[\hat{I}_0+\hat{S}_0-S_\infty(0)\right]>0, \qquad \hat{I}_0+\hat{S}_0-2S_\infty(0)<0 \quad \text{(B10)}$$

$$1-\mathcal{R}_0\left[\hat{I}_0+\hat{S}_0-S_\infty(0)\right]<0, \qquad \hat{I}_0+\hat{S}_0-2S_\infty(0)>0 \quad \text{(B11)}$$

Suppose (B10) holds. Then (B9) is strictly less than

$$-S_\infty(0)\left[2(\hat{I}_0+\hat{S}_0)-S_\infty(0)\right]-S_\infty(0)\left\{1-\mathcal{R}_0\left[\hat{I}_0+\hat{S}_0-S_\infty(0)\right]\right\}\left[\hat{I}_0+\hat{S}_0-2S_\infty(0)\right] \quad \text{(B12)}$$

$$<-S_\infty(0)\left[2(\hat{I}_0+\hat{S}_0)-S_\infty(0)\right]-S_\infty(0)\left[\hat{I}_0+\hat{S}_0-2S_\infty(0)\right] \quad \text{(B13)}$$

$$=-3S_\infty(0)\left[\hat{I}_0+\hat{S}_0-S_\infty(0)\right]. \quad \text{(B14)}$$

Equation (B12) follows from eliminating the first two negative terms of (B9). Equation (B13) follows from substituting 1, which is greater than the factor in braces, for the factor in braces. The fact that this substitution results in an increase in (B13) follows from (B10). Straightforward algebra yields (B14), which is negative by familiar arguments.

Suppose (B11) holds. Then (B9) is strictly less than

$$S_\infty(0)\left[1-\mathcal{R}_0(\hat{I}_0+\hat{S}_0)\right]-S_\infty(0)\left\{1-\mathcal{R}_0\left[\hat{I}_0+\hat{S}_0-S_\infty(0)\right]\right\}\left[\hat{I}_0+\hat{S}_0-2S_\infty(0)\right] \quad \text{(B15)}$$

$$<S_\infty(0)\left[1-\mathcal{R}_0(\hat{I}_0+\hat{S}_0)\right]-S_\infty(0)\left\{1-\mathcal{R}_0\left[\hat{I}_0+\hat{S}_0-S_\infty(0)\right]\right\} \quad \text{(B16)}$$

$$=-\mathcal{R}_0 S_\infty(0)^2. \quad \text{(B17)}$$

Equation (B15) follows from eliminating the second and third two negative terms from (B9). Equation (B16) follows from substituting 1 for the last factor, $\hat{I}_0+\hat{S}_0-2S_\infty(0)$. To see that this increases the expression, note $\hat{I}_0+\hat{S}_0-2S_\infty(0)<\hat{I}_0+\hat{S}_0\leq 1$, where the last inequality holds since the size of the infected and susceptible subpopulations at date 0, $\hat{I}_0+\hat{S}_0$, cannot exceed the size of the entire population, normalized to 1. The fact that substituting 1 for $\hat{I}_0+\hat{S}_0-2S_\infty(0)$ increases (B15) follows from (B11). Straightforward algebra yields (B17), which is obviously negative.

In sum, we have shown (B9) is negative for $\mathcal{R}_0<\mathcal{R}_0'$, implying (B7) is concave. In the limit $\mathcal{R}_0\downarrow 0$, (B7) approaches $\hat{S}_0-S_\infty(0)$, which is positive by Lemma 4. These facts are sufficient to establish that MEX_c^* is quasiconcave in case (SR1) by the same arguments used for MSB_c^* above. These are all the facts needed to prove that MEX_c^* has a unique restricted local maximum over $\mathcal{R}_0\leq\mathcal{R}_0''$, which is a global maximum on that restricted set.

We next investigate the behavior of MEX_c^* in (SR3) and (SR4). Calculations similar to those used above can be used to show a function determining the sign of $\partial MEX_c^*/\partial \mathcal{R}_0$ in cases (SR3) and (SR4),

$$\left[(1-\theta)\hat{S}_0 - S_\infty(0)\right]\left\{1-\mathcal{R}_0[\hat{I}_0+(1-\theta)\hat{S}_0]\right\}+\mathcal{R}_0 S_\infty(\hat{S}_0)\left[1-\mathcal{R}_0 S_\infty(\hat{S}_0)\right]\left[\hat{I}_0+(1-\theta)\hat{S}_0-S_\infty(\hat{S}_0)\right]. \quad \text{(B18)}$$

is concave. Thus, (B18) has at most two roots in cases (SR3) and (SR4), at most one of which is a local maximum for MEX_c^*. Taking the limit of the table entry for MEX_c^* in cases (SR3) and (SR4) and substituting the limit $\lim_{\mathcal{R}_0\uparrow\infty} S_\infty(\hat{S}_0) = \lim_{\mathcal{R}_0\uparrow\infty}[\mathcal{R}_0 S_\infty(\hat{S}_0)] = 0$ by Lemma 8 yields $\lim_{\mathcal{R}_0\uparrow\infty} MEX_c^* = 0$. Since the limit $\mathcal{R}_0 \uparrow \infty$ produces an infimum for MEX_c^*, not a supremum, the restricted supremum of MEX_c^* over $\mathcal{R}_0 > \mathcal{R}_0''$ is either the lower boundary of case (SR3), i.e., $\mathcal{R}_0''$, or the interior restricted maximum. If $\mathcal{R}_0''$ provides the restricted supremum over $\mathcal{R}_0 > \mathcal{R}_0''$, this cannot be a global maximum since MSB_c^* is decreasing in case (SR2); the restricted maximum over $\mathcal{R}_0 \leq \mathcal{R}_0''$ must then be the global maximum. *Q.E.D.*

Completing Proof of Proposition 2

It remains to provide results for $R_\infty(Q_m^*)$ and MEX_m^*, omitted from the proof in Appendix A.

Results for Π_m^*: The result is a consequence of the Envelope Theorem. Monopoly profit can be written

$$\Pi_m^* = \theta H\left[1-\frac{S_\infty(Q_m^*)}{S_0(Q_m^*)}-\tilde{c}\right]Q_m^*. \quad \text{(B19)}$$

This is a function of $\mathcal{R}_0$ indirectly through its dependence on Q_m^*, which in turns depends on $\mathcal{R}_0$. It also depends on $\mathcal{R}_0$ because $S(Q)$ is a function of $\mathcal{R}_0$ (although the argument is omitted for brevity). If Q_m^* is an interior solution, as in case (SR2) and (SR3), the first-order condition ensures that the indirect effect of $\mathcal{R}_0$ on Π_m^* through Q_m^* equals 0. Only the direct effect remains. Hence,

$$\frac{\partial \Pi_m^*}{\partial \mathcal{R}_0} = \left[\frac{-\theta H Q_m^*}{S_0(Q_m^*)}\right]\frac{\partial S_\infty(Q_m^*)}{\partial \mathcal{R}_0}, \quad \text{(B20)}$$

which is positive since the derivative on the right-hand side is negative by (A20).

Results for $R_\infty(Q_m^*)$: Since $R_\infty(Q_m^*) = R_\infty(Q_c^*)$ for all $\mathcal{R}_0$ in case (SR1),

$$\lim_{\mathcal{R}_0\downarrow\mathcal{R}_0'} R_\infty(Q_m^*) = \lim_{\mathcal{R}_0\downarrow\mathcal{R}_0'} R_\infty(Q_c^*) = 1-(1-\tilde{c})\hat{S}_0, \quad \text{(B21)}$$

where the first equality follows by continuity since $\mathcal{R}_0'$ is the upper bound on case (SR1) by (A18) and the second equality follows from (B3).

We proceed to compare (B21) to the limits of $R_\infty(Q_m^*)$ for extreme values of $\mathcal{R}_0$. We have

$$\lim_{\mathcal{R}_0\downarrow 0} R_\infty(Q_m^*) = \lim_{\mathcal{R}_0\downarrow 0} R_\infty(Q_c^*) = 1-\lim_{\mathcal{R}_0\downarrow 0} S_\infty(0) = 1-\hat{S}_0, \quad \text{(B22)}$$

where the first equality follows since $R_\infty(Q_m^*) = R_\infty(Q_c^*)$ for all $\mathcal{R}_0$ in case (SR1), the second equality follows from the entry for $R_\infty(Q_c^*)$ in case (SR1) in Table 1, and the third equality follows from Lemma 7. Equation (B22) is less than (B21). At the other extreme,

$$\lim_{\mathcal{R}_0\uparrow\infty} R_\infty(Q_m^*) = \lim_{\mathcal{R}_0\uparrow\infty} R_\infty(Q_c^*) = 1-\theta\hat{S}_0, \quad \text{(B23)}$$

where the first equality follows by continuity since $R_\infty(Q_m^*) = R_\infty(Q_c^*)$ for all $\mathcal{R}_0$ in case (SR4), the second equality follows from the entry for $R_\infty(Q_c^*)$ in cases (SR3) and (SR4) in Table 1, and the third equality follows from (B5). Equation (B23) is weakly less than (B21) if $\tilde{c} \geq 1-\theta$. Since $R_\infty(Q_m^*)$ is greater at the interior $\mathcal{R}_0'$ than at extreme values of $\mathcal{R}_0$, $R_\infty(Q_m^*)$ must have an interior maximum.

Results for MEX_m^*: Arguments similar to those used in Appendix A for MSB_m^* the preceding can be used to show $\lim_{\mathcal{R}_0 \downarrow 0} MEX_m^* = \lim_{\mathcal{R}_0 \uparrow \infty} MEX_c^* = 0$. Hence, MEX_m^* is higher for interior values of $\mathcal{R}_0$ than extreme values and thus attains an interior maximum. *Q.E.D.*

Appendix B2. Cournot Competition

This appendix extends the analysis to Cournot competition, which nests the perfectly competitive and monopoly market structures studied in the text. Under Cournot competition, the vaccine is manufactured by $n \geq 1$ homogeneous firms, which choose quantities each period simultaneously.

Case (SR1) from Table 1, which involved no sales under perfect competition, will also involve no sales under Cournot since firms mark up marginal costs. Thus the entries in case (SR1) from both Tables 1 and 2 will also apply to Cournot.

For the remainder of this section, suppose $\mathcal{R}_0 > |\ln(1-\tilde{c})|/(\hat{I}_0 + \tilde{c}\hat{S}_0)$. As in the long-run analysis, firm i's profit equals $[P(q_i+Q_{-i})-c]q_i = [MPB(q_i+Q_{-i})-c]q_i$. Taking the first-order condition with respect to q_i and then imposing symmetry by substituting $q_i^* = Q^*/n$ yields an equation for the interior equilibrium market output analogous to (27) for a monopoly:

$$MPB(Q_n^*) = P(Q_n^*) = c \Bigg/ \left\{ 1 - \frac{\theta Q_n^* \mathcal{R}_0 S_\infty(Q_n^*)}{n S_0(Q_n^*)[1-\mathcal{R}_0 S_\infty(Q_n^*)]} \right\}. \tag{B24}$$

The only difference is the appearance of n in the denominator of the factor in curly braces. One can see this nests monopoly, setting $n = 1$, and perfect competition, taking the limit $n \uparrow \infty$.

This interior solution is the equilibrium market output under Cournot if $Q_n^* < \hat{S}_0$. Otherwise, $Q_n^* = \hat{S}_0$, and all firms produce an equal share $q_n^* = \hat{S}_0/n$ in the symmetric equilibrium.

Appendix B3. Consumer Heterogeneity

The model in the text assumes consumers are homogeneous. This appendix introduces consumer heterogeneity and shows that the key result regarding the nonmonotonicities of the marginal externality continues to hold in this extension.

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.(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.) 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. 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.

With homogeneous consumers, we showed the marginal private benefit can be written $MPB(Q) = \theta H \Phi_I(Q)$, the product of efficacy, harm, and probability of contracting the disease. With consumer heterogeneity, consumer i's marginal private benefit becomes $MPB_i(Q) = \theta H_i \Phi_I(Q)$.

Incorporating heterogeneity in some of the normative measures requires additional work to keep track of the high-value consumers who end up purchasing. We have

$$SB(Q) = \left\{ \left[1-\Phi_I(Q)\right] \int_0^{\hat{H}} H_i f(H_i) dH_i + \left[1-\Phi_I(Q)-\theta\Phi_I(Q)\right] \int_{\hat{H}}^{\infty} H_i f(H_i) dH_i \right\} \hat{S}_0. \tag{B25}$$

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 $1-\Phi_I(Q)$. The second integral reflects the expected health experienced by those who purchase. Consumer i in this group obtains H_i if either they would not have been infected anyway (probability $1-\Phi_I(Q)$) or would have been infected without a vaccine but receive the vaccine protection (probability $\theta\Phi_I(Q)$). The final factor $\hat{S}_0$ allows the per-consumer surplus given by the integrals to be scaled up to the population of potential consumers. Differentiating (B25) yields

$$MSB(Q) = \left\{ -\frac{\partial \Phi_I(Q)}{\partial Q} \left[E(H_i) - \theta \int_{\hat{H}}^{\infty} H_i f(H_i) dH_i \right] + \theta \Phi_I(Q) \hat{H} f(\hat{H}) \frac{\partial \hat{H}}{\partial Q} \right\} \hat{S}_0. \tag{B26}$$

To compute $\partial \hat{H}/\partial Q$, note threshold consumer type $\hat{H}$ is given as an implicit function of Q by $Q = \bar{F}(\hat{H})\hat{S}_0$. Totally differentiating this identity with respect to Q and rearranging yields $\partial \hat{H}/\partial Q = 1/f(\hat{H})\hat{S}_0$. Substituting this derivative into (B26) shows that the last term equals $\theta \hat{H} \Phi_I(Q)$. This is the private benefit of the threshold consumer, equal to MPB^* when evaluated at the equilibrium Q^*. Subtracting to compute $MEX^* = MSB^* - MPB^*$ leaves just the first term of (B26), as stated in the following lemma.

Lemma B1. *In the model with heterogeneity in consumer harm H_i, the marginal externality equals*

$$MEX^* = -\frac{\partial \Phi_I(Q^*)}{\partial Q} \left[E(H_i) - \theta \int_{\hat{H}(Q^*)}^{\infty} H_i f(H_i) dH_i \right] \hat{S}_0. \tag{B27}$$

Intuitively, Lemma B1 says that the marginal externality is proportional to $-\partial \Phi_I(Q^*)/\partial Q$, 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 $\mathcal{R}_0 \downarrow 0$ since a noninfectious disease presents no danger of infection. The factor also approaches 0 as $\mathcal{R}_0 \uparrow \infty$ 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 $\mathcal{R}_0$. Thus, MEX^* approaches 0 for extreme values of $\mathcal{R}_0$, implying it is nonmonotonic in $\mathcal{R}_0$, as the following proposition states.

Proposition B1. *In the model with heterogeneity in consumer harm H_i, MEX* is nonmonotonic under both perfect competition and monopoly.*

Proof. It remains to analyze the limits of $\partial\Phi_I(Q^*)/\partial Q$ as $\mathcal{R}_0 \downarrow 0$ and $\mathcal{R}_0 \uparrow \infty$, showing that the limits equal 0 for both market structures. Differentiating (15) and substituting from (14) yields

$$\frac{\partial\Phi_I(Q)}{\partial Q} = \frac{-\theta\Phi_I(Q)\mathcal{R}_0 S_\infty(Q)}{S_0(Q)[1-\mathcal{R}_0 S_\infty(Q)]}. \tag{B28}$$

Lemma 7 states $\lim_{\mathcal{R}_0\downarrow 0}[\mathcal{R}_0 S_\infty(Q)] = 0$, implying $\lim_{\mathcal{R}_0\downarrow 0}\partial\Phi_I(Q)/\partial Q = 0$ by (B28). Lemma 8 states $\lim_{\mathcal{R}_0\uparrow\infty}[\mathcal{R}_0 S_\infty(Q)] = 0$, implying $\lim_{\mathcal{R}_0\uparrow\infty}\partial\Phi_I(Q)/\partial Q = 0$ by (B28). These limits both hold for all Q, including $Q = Q_c^*$ and $Q = Q_m^*$. *Q.E.D.*

Appendix B4. Long-Run Analysis of Endemic Disease

In this appendix, we move from a short-run analysis of an epidemic to the analysis of an endemic disease disease such as HIV that persists in the population over the long run. We begin by making the necessary modifications of the SIR model to capture an endemic disease. Reflecting our long-run perspective, we focus on the steady-state equilibrium.

Epidemiological Model with Demographic Flows

For the SIR model to represent the long-run market for the vaccine, two main modifications are needed. First, demographic flows (births and deaths) need to be included. Let $\mu \in (0,1)$ denote the mortality rate from causes other than the disease, referred to as "natural causes." Restricting attention to a non-fatal disease as before and setting the birth rate to equal to μ leaves the population mass constant over time. We continue to to normalize the total population mass to 1, as indicated by equation (1). The second modification is that, rather than a one-shot campaign, the vaccine is administered to the population in a continuous flow. Reinterpret $Q \geq 0$ as 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. While Q could vary over time in principle, we omit a time subscript anticipating the later solution for its equilibrium value in the steady state. See Martcheva (2015) for a textbook treatment of the enhanced version of an SIR model with vaccination and demographic flows used here.

The following equations describe the evolution of consumer compartments in the long-run version of the model:

$$\dot{S}_t = \mu - \theta Q - \beta I_t S_t - \mu S_t \tag{B29}$$

$$\dot{I}_t = \beta I_t S_t - \alpha I_t - \mu I_t \tag{B30}$$

$$\dot{R}_t = \alpha I_t - \mu R_t \tag{B31}$$

$$\dot{V}_t = \theta Q - \mu V_t. \tag{B32}$$

Equations (B29)–(B32) are similar to (2)–(5). The last term subtracted from each equation reflects the flow out of each compartment due to mortality. Consistent with our assumption that the disease

is non-fatal, the mortality rate is the same μ for infected consumers as for other compartments. Each instant, μ newborns enter the market as susceptibles, captured by the first term added to the right-hand side of (B29). We assume vaccines are administered only to newborns. (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.) The Q vaccines each administered each instant result in θQ successful vaccinations and thus θQ consumers flowing out of compartment S_t and into compartment V_t.

As before, we will work with the basic reproductive ratio $\mathcal{R}_0$ rather than the transmission rate β. With demographic flows, the expression for $\mathcal{R}_0$ needs to be modified from equation (11):

$$\mathcal{R}_0 = \frac{\beta}{\alpha+\mu}. \tag{B33}$$

The reciproal of the infected spell given by the new denominator of $\mathcal{R}_0$ reflects the fact that mortality is now a competing risk (along with recovery) for an individual to exit the infected state. To derive the infected spell, note that recovery has hazard $\lambda_R(t) = \alpha$, and mortality has hazard $\lambda_M(t) = \mu$. The combined hazard of exiting the infected state is $\lambda_{EI}(t) = \lambda_R(t) + \lambda_M(t)$. In a Poisson duration model, the duration of a spell equals the reciprocal of the hazard, implying $1/\lambda_{EI}(t) = 1/(\alpha+\mu)$.

Endemic Steady State

As in the text, we append Q as an argument to equilibrium values to emphasize their dependence on that key variable to be endogenized later and use $S_\infty(Q)$, $I_\infty(Q)$, $R_\infty(Q)$, and $V_\infty(Q)$ to denote limiting compartment values. Limiting values are now interpreted as steady-state values for an endemic disease rather than values converged to at the end of an epidemic.

Steady-state compartment values can be found by solving the system of equations formed by setting $\dot{S}_t = \dot{I}_t = \dot{R}_t = \dot{V}_t = 0$ in equations (B29)–(B32). The unique stable solution is

$$S_\infty(Q) = \min\left(1 - \frac{\theta}{\mu}Q, \frac{1}{\mathcal{R}_0}\right) \tag{B34}$$

$$I_\infty(Q) = \max\left[0, \frac{\mu}{\alpha+\mu}\left(1 - \frac{1}{\mathcal{R}_0} - \frac{\theta}{\mu}Q\right)\right] \tag{B35}$$

$$R_\infty(Q) = \frac{\alpha}{\mu} I_\infty(Q) \tag{B36}$$

$$V_\infty(Q) = \frac{\theta}{\mu}Q. \tag{B37}$$

A trivial solution involving $I_\infty(Q) = 0$ always exists, but it is unstable when $\mathcal{R}_0 > \mu/(\mu - \theta Q)$.

By equation (B35), if $\mathcal{R}_0 \leq 1$, then $I_\infty(Q) = 0$ for all $Q \geq 0$. The disease dies out in the steady state with or without a vaccine. For a non-trivial vaccine market to exist, $\mathcal{R}_0 > 1$. In that case, whether the disease dies out in the steady state depends on Q: $I_\infty(Q) = 0$ if and only if $Q \geq (\mu/\theta)(1 - 1/\mathcal{R}_0)$. Combining the various cases in a way that will be convenient for the subsequent analysis, we can rewrite

$$I_\infty(Q) = \begin{cases} 0 & Q \geq \mathcal{Q}_0 \\ \dfrac{\mu}{\alpha+\mu}\left(1 - \dfrac{1}{\mathcal{R}_0} - \dfrac{\theta}{\mu}Q\right) & Q < \mathcal{Q}_0, \end{cases} \tag{B38}$$

where

$$\mathcal{Q}_0 = \max\left[0, \frac{\mu}{\theta}\left(1 - \frac{1}{\mathcal{R}_0}\right)\right] \tag{B39}$$

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

Consumer Demand

We next turn to economic features of the vaccine market, starting with demand. Continue to assume that consumers are risk neutral. Return for now to the assumption that consumers are homogeneous; the case of heterogeneous consumers is analyzed in Online Appendix B. To avoid computing transition paths, focus on the steady-state in the limiting case without discounting.

As before, a susceptible consumer demands the vaccine if his or her marginal private benefit $MPB(Q) = \theta H \Phi_I(Q)$ exceeds the vaccine's price, P. The modifications to the epidemiological model to suit an endemic disease require some work to derive a new expression for $\Phi_I(Q)$. A susceptible consumer who has chosen not to be vaccinated faces two remaining competing risks to exit that state: infection, which has hazard $\lambda_I(t,Q) = \beta I_\infty(Q)$ each instant, and mortality from natural causes, which has hazard $\lambda_M(t) = \mu$. The combined hazard of exiting the susceptible state is $\lambda_{ES}(t,Q) = \lambda_I(t,Q) + \lambda_M(t)$. By standard results for competing Poisson risks, the consumer's cumulative risk of exiting the susceptible state by age t is $\Lambda_{ES}(t,Q) = \int_0^t \lambda_{ES}(\tau,Q)d\tau$, probability of surviving as susceptible to age t is $e^{-\Lambda_{ES}(t,Q)}$, and likelihood of exit due to infection is $\phi_I(t,Q) = \lambda_I(t,Q)e^{-\Lambda_{ES}(t,Q)}$. To compute the probability that the consumer experiences an infection at some point over his or her lifetime, we integrate this cause-specific likelihood:

$$\Phi_I(Q) = \int_0^\infty \phi_I(t,Q)dt = \frac{\beta I_\infty(Q)}{\beta I_\infty(Q) + \mu} = \frac{(\alpha+\mu)\mathcal{R}_0 I_\infty(Q)}{(\alpha+\mu)\mathcal{R}_0 I_\infty(Q) + \mu}. \tag{B40}$$

Substituting for $I_\infty(Q)$ from (B38) yields

$$\Phi_I(Q) = \begin{cases} 0 & Q \geq \mathcal{Q}_0 \\ 1 - \dfrac{1}{(1-\theta Q/\mu)\mathcal{R}_0} & Q < \mathcal{Q}_0. \end{cases} \tag{B41}$$

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

$$MPB(Q) = \begin{cases} 0 & Q \geq \mathcal{Q}_0 \\ \theta H\left[1 - \dfrac{1}{(1-\theta Q/\mu)\mathcal{R}_0}\right] & Q < \mathcal{Q}_0. \end{cases} \tag{B42}$$

Applying the same steps used in Section 2.3 to derive demand from a comparison of $MPB(Q)$ to P using the new expression for $MPB(Q)$ in (B42) yields demand curve

$$D(P) = \begin{cases} 0 & P > MPB(0) \\ d(P) & P \in [MPB(\mu), MPB(0)] \\ \mu & P < MPB(\mu), \end{cases} \tag{B43}$$

where now

$$d(P) = \frac{\mu}{\theta}\left[1 - \frac{1}{(1-P/\theta H)\mathcal{R}_0}\right] \tag{B44}$$

is the fraction of consumers who buy when they are indifferent, pinned down by setting (B42) equal to P and inverting.

Demand considerations alone allow us to establish that an important result due to Geoffard and Philipson (1997)—that 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.

Proposition B2. *Suppose that the disease is eradicated in steady-state equilibrium of a market with or without government subsidies, i.e., $I_\infty^* = 0$. Then either the vaccine is free to consumers ($P^* = 0$) or no vaccine is purchased ($Q^* = 0$).*

Proof. Suppose $I_\infty(Q^*) = 0$ and $Q^* > 0$. We will show $P^* = 0$. Substituting $I_\infty(Q^*) = 0$ into (B35) and rearranging yields $Q^* \geq (\mu/\theta)(1-1/\mathcal{R}_0)$. This inequality together with $Q^* > 0$ implies $Q^* \geq \mathcal{Q}_0$ by (B39), implying $MPB^* = 0$ by (B42). Since $MPB(Q)$ is weakly decreasing in Q, $MPB(\mu) \leq MPB^* = 0$. Then $Q^* = D(P^*)$ cannot be determined by the third branch of equation (B43); if it were, $P^* < MPB(\mu) = 0$, violating the nonnegativity of prices. Since $Q^* > 0$, Q^* cannot be determined by the first branch of (B43) either. Therefore,

$$Q^* = d(P^*) = \frac{\mu}{\theta}\left[1 - \frac{1}{(1-P/\theta H)\mathcal{R}_0}\right]. \tag{B45}$$

Since $0 = MPB^* = MPB(Q^*)$, we have

$$Q^* \geq \mathcal{Q}_0 \geq \frac{\mu}{\theta}\left(1 - \frac{1}{\mathcal{R}_0}\right), \tag{B46}$$

where the first inequality follows from equation (B42) and the second from (B39). Combining (B45) with (B46) and rearranging yields $P^* \leq 0$, implying $P^* = 0$ by the nonnegativity of prices. *Q.E.D.*

To gain some intuition for the proof, if the disease is eradicated in steady-state equilibrium, then $\Phi_I(Q) = 0$, implying $MPB(Q) = 0$. But consumers with no marginal private benefit will not purchase the vaccine at a positive price.

Firm Supply

Firm behavior is the same as that assumed in the short-run analysis. Firms produce at constant marginal and average cost $c \in (0, \theta H)$ per vaccine course. Perfectly competitive firms supply the vaccine at price c. A monopoly sets a price maximizing industry profit, Π, now measured as a flow each instant in the steady state.

Normative Measures

In the short-run analysis, the fixed population allowed us to measure social benefit in terms of a stock, in particular, the stock of avoided harm for those never infected. With demographic flows, the stock of avoided harm is undefined (infinite), so a flow measure is needed. Define the total social benefit flowing each instant from the vaccine in the steady state as $SB(Q) = h[1-I_\infty(Q)]$, the product of the flow health benefit h to an individual and the number of healthy individuals $1-I_\infty(Q)$ at any

instant in steady-state equilibrium. We earlier defined H as the expected harm to an individual over the whole disease spell. Using the result from Section **??** that the expected disease spell equals $1/(\alpha+\mu)$ yields $H = h/(\alpha+\mu)$, or $h = (\alpha+\mu)H$. Substituting this as well as equation (B35) into the definition of social benefit yields

$$SB(Q) = \begin{cases} (\alpha+\mu)H & Q \geq \mathcal{Q}_0 \\ (\alpha+\theta Q+\mu/\mathcal{R}_0)H & Q \leq \mathcal{Q}_0. \end{cases} \tag{B47}$$

Marginal social benefit is the derivative of (B47). Technically, the derivative $\partial SB(\mathcal{Q}_0)\partial Q$ does not exist, but the left and right derivatives, respectively $\partial SB(\mathcal{Q}_0)/\partial Q^-$ and $\partial SB(\mathcal{Q}_0)/\partial Q^+$, do. Setting $MSB(\mathcal{Q}_0) = \partial SB(\mathcal{Q}_0)/\partial Q^+$ yields

$$MSB(Q) = \begin{cases} 0 & Q \geq \mathcal{Q}_0 \\ \theta H & Q < \mathcal{Q}_0. \end{cases} \tag{B48}$$

Subtracting (B47) from (B42) the expression for the marginal externality

$$MEX(Q) = \begin{cases} 0 & Q \geq \mathcal{Q}_0 \\ \dfrac{\theta H}{(1-\theta Q/\mu)\mathcal{R}_0} & Q < \mathcal{Q}_0. \end{cases} \tag{B49}$$

For $Q \geq \mathcal{Q}_0$, enough vaccine is available to eradicate the disease in the steady state, eliminating the marginal benefit of vaccine. Hence, $MPB(Q) = MSB(Q) = MEX(Q) = 0$ for $Q \geq \mathcal{Q}_0$.

As in equation (22), social welfare continues to be given by $W(Q) = SB(Q) - cQ$, now using equation (B47) for $SB(Q)$. To characterize the first-best quantity Q^{**} maximizing maximizes $W(Q)$, under maintained assumption (20), we can see from (B48) that $Q^{**} = \mathcal{Q}_0$ unless this entails vaccinating more than the flow μ of newborns each instant, in which case $Q^{**} = \mu$. Thus,

$$Q^{**} = \min(\mathcal{Q}_0, \mu). \tag{B50}$$

Perfectly Competitive Equilibrium

Under perfect competition $P_c^* = c$. The remaining equilibrium variables can be computed using straightforward algebra applied to the supplied formulas. The computations are easier than in the short-run analysis since compartments did not have closed-form expressions there but do here. Table B1 reports the steady-state equilibrium values of selected variables under perfect competition for various intervals of $\mathcal{R}_0$. We proceed to formally verify the table entries next and then provide intuition for the entries.

As a preliminary step in verifying the table entries, we will verify that the intervals in the column headings are ordered as given. The fact that $c > 0$ implies $\tilde{c} > 0$. Assumption (20) implies $\tilde{c} < 1$. Hence $\tilde{c} \in (0,1)$, implying $1/(1-\tilde{c}) > 1$, implying the case (LR2) interval for $\mathcal{R}_0$ is to the right of the case (LR1) interval. The ordering of the cases is obvious if $\theta < 1$. If $\theta = 1$, then cases (LR4) and (LR5) fail to exist.

Turning to the individual cases, in case (LR1), $\mathcal{R}_0 \leq 1$ implies $\mathcal{Q}_0 = 0$ by equation (B39). Hence, $MPB(Q) = MSB(Q) = MEX(Q) = 0$ for all $Q \geq 0$ by equations (B42), (B48), and (B49). The fact that $MPB(Q) = 0$ for all $Q \geq 0$ implies $Q_c^* = D(c) = 0$ for $c > 0$ by (B43). Equilibrium welfare is $W_c^* = SB_c^* - cQ_c^* = (\alpha+\mu)H$ by (B47).

Table B1: Long-Run Equilibrium Variables under Perfect Competition as Functions of $\mathcal{R}_0$

	Case			
	(LR1)	(LR2)	(LR3)	(LR4), (LR5)
Variable	$\mathcal{R}_0 \in [0,1]$	$\mathcal{R}_0 \in \left(1, \frac{1}{1-\tilde{c}}\right]$	$\mathcal{R}_0 \in \left(\frac{1}{1-\tilde{c}}, \frac{1}{(1-\theta)(1-\tilde{c})}\right]$	$\mathcal{R}_0 \in \left(\frac{1}{(1-\theta)(1-\tilde{c})}, \infty\right)$
P_c^*	c	c	c	c
Q_c^*	0	0	$\frac{\mu}{\theta}\left[1-\frac{1}{(1-\tilde{c})\mathcal{R}_0}\right]$	μ
Π_c^*	0	0	0	0
$I_\infty(Q_c^*)$	0	$\frac{\mu}{\alpha+\mu}\left(1-\frac{1}{\mathcal{R}_0}\right)$	$\frac{\mu}{\alpha+\mu}\left[\frac{\tilde{c}}{(1-\tilde{c})\mathcal{R}_0}\right]$	$\frac{\mu}{\alpha+\mu}\left(1-\theta-\frac{1}{\mathcal{R}_0}\right)$
MPB_c^*	0	$\theta H\left(1-\frac{1}{\mathcal{R}_0}\right)$	c	$\theta H\left[1-\frac{1}{(1-\theta)\mathcal{R}_0}\right]$
MSB_c^*	0	θH	θH	θH
MEX_c^*	0	$\frac{\theta H}{\mathcal{R}_0}$	$\theta H(1-\tilde{c})$	$\frac{\theta H}{(1-\theta)\mathcal{R}_0}$
W_c^*	$H(\alpha+\mu)$	$H\left(\alpha+\frac{\mu}{\mathcal{R}_0}\right)$	$H\left[\alpha+\mu(1-\tilde{c})\right]$	$H\left[\alpha+\theta\mu(1-\tilde{c})+\frac{\mu}{\mathcal{R}_0}\right]$

Notes: The distinction between cases (LR4) and (LR5) in the last column, relevant for monopoly in the next table, is irrelevant for perfect competition here.

In case (LR2), $\mathcal{R}_0 > 1$ implies $\mathcal{Q}_0 > 0$ by (B39). Thus, by (B42),

$$MPB(0) = \theta H(1-1/\mathcal{R}_0) < \theta H\tilde{c} = c, \tag{B51}$$

where the inequality follows from $\mathcal{R}_0 < 1/(1-\tilde{c})$. We can verify the claim in the text that $MPB(Q)$ is weakly decreasing in Q by differentiating (B42):

$$-\frac{\theta^2 H}{\mu\mathcal{R}_0}\left(1-\frac{\theta Q}{\mu}\right)^{-2} < 0, \tag{B52}$$

implying that $MPB(Q)$ is weakly decreasing in Q. Therefore, $MPB(Q) \leq MPB(0) < c$, implying $Q_c^* = D(c) = 0$ by (B43).

Skipping over case (LR3) to cases (LR4) and (LR5), assume $\theta < 1$, so that these cases exist. We have

$$\mathcal{R}_0 > 1/(1-\theta)(1-\tilde{c}) > 1/(1-\tilde{c}), \tag{B53}$$

implying $\mathcal{Q}_0 > \mu$ by (B39). Since $Q \leq \mu < \mathcal{Q}_0$, $MPB(Q)$ is given by the second branch in (B42), implying

$$MPB(Q) \geq MPB(\mu) \geq \theta H\left[1-\frac{1}{(1-\theta)\mathcal{R}_0}\right] \geq \theta H[1-(1-\tilde{c})] = c, \tag{B54}$$

where the first step follows from the fact that $MPB(Q)$ is weakly decreasing, the second from substituting $Q = \mu$ into (B42), the third step from (B53), and the last from of $\tilde{c} = c/\theta H$. Thus, $Q_c^* = \mu$

by (B43). Since $Q < \mathcal{Q}_0$, the relevant branch in equations (B47)–(B49) for $SB(Q)$, $MSB(Q)$, and $MEX(Q)$ is the second. Substituting $Q_c^* = \mu$ into these equations gives SB_c^*, MSB_c^*, and MEX_c^*. Substituting Q_c^* and SB_c^* into the welfare formula gives W_c^*.

In case (LR3), $\mathcal{R}_0 \geq 1/(1-\tilde{c})$ implies $MPB(0) \geq c$ by (B51). Further, $\mathcal{R}_0 \leq 1/(1-\theta)(1-\tilde{c})$ implies $MPB(\mu) \leq c$ by (B54). Hence, by (B43), $Q_c^* = d(c)$. Substituting c into (B44) yields the table entry for Q_c^*. Using (B39), one can show $Q_c^* < \mathcal{Q}_0$, implying that the relevant branch in equations (B47)–(B49) for $SB(Q)$, $MSB(Q)$, and $MEX(Q)$ is the second. The equilibrium value of the remaining variables can be derived as in the previous paragraph.

The final step in verifying the entries to Table B1 is to note that since $P_c^* = c$ in all cases, $\Pi_c^* = 0$ in all cases.

To gain some intuition for the entries to Table B1, note that for $\mathcal{R}_0 \leq 1$, labeled case (LR1), infectiveness is so low 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 $\mathcal{R}_0$ just above 1, the disease is infective enough not to disappear in the steady state but not infective enough to justify its purchase at the competitive price. In this case, labeled (LR2), though the vaccine has a positive marginal social and private benefit, the marginal private benefit is below c even if no other consumer purchases. For $\mathcal{R}_0$ in the next interval, labeled (LR3), some but not all newborns purchase. For extreme values of $\mathcal{R}_0$, the disease can conceivably be so infective that all μ newborns purchase at the competitive price. The existence of this case, labeled (LR4), 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. Since all consumers purchase in this case, there is no underconsumption distortion; the first best is obtained.

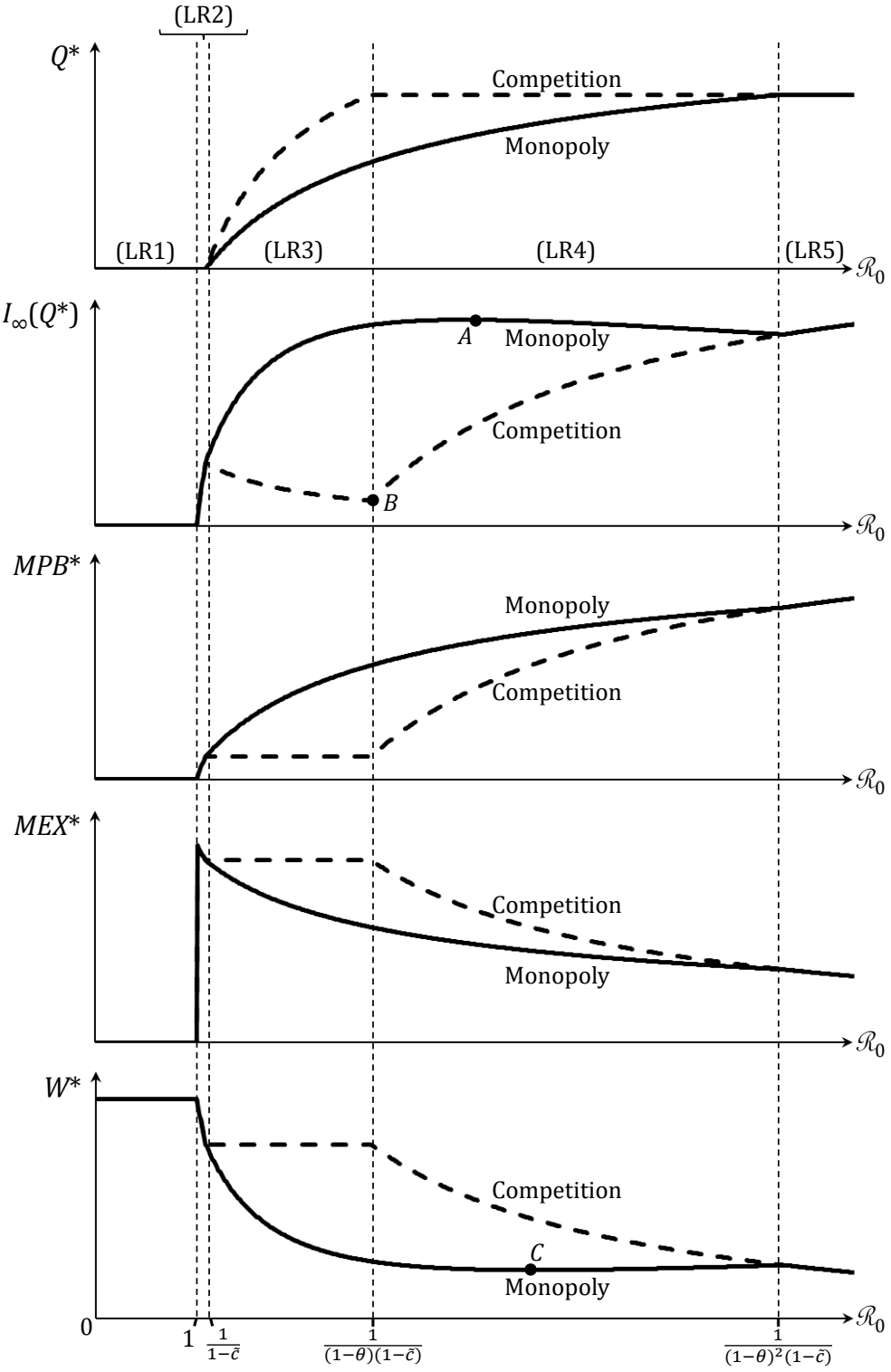
To help visualize how the variables in the table vary with $\mathcal{R}_0$, Figure B1 graphs a further selection of them.

The next proposition distills the comparative-static effects of an increase in $\mathcal{R}_0$ on the steady-state equilibrium under perfect competition from the entries in Table B1. Among other notable results, the nonmonotonic behavior of infections and the marginal externality uncovered in the short-run analysis persists in the long-run analysis.

Proposition B3. *Consider the comparative-static effect of $\mathcal{R}_0$ on steady-state equilibrium under perfect competition.*

- *Price and industry profit are constant for $\mathcal{R}_0 > 0$, with $P_c^* = c$ and $\Pi_c^* = 0$.*
- *Q_c^* and MPB_c^* are weakly increasing in $\mathcal{R}_0$ for all $\mathcal{R}_0 > 0$.*
- *$I_\infty(Q_c^*)$ is nonmonotonic, reaching a local maximum $I_\infty(Q_c^*) = \mu\tilde{c}/(\alpha+\mu)$ at $\mathcal{R}_0 = 1/(1-\tilde{c})$, which is a global maximum if $\tilde{c} > 1-\theta$.*
- *Marginal social benefit is constant for all $\mathcal{R}_0 > 1$, with $MSB_c^* = \theta H$.*
- *The marginal externality is nonmonotonic over $\mathcal{R}_0 > 0$, approaching the supremum $MEX_c^* = \theta H$ as $\mathcal{R}_0 \downarrow 1$.*
- *W_c^* is weakly decreasing in $\mathcal{R}_0$ for all $\mathcal{R}_0 > 0$.*

Figure B1: Graphs of Long-Run Equilibrium Variables as Functions of $\mathcal{R}_0$



Notes: Graph of formulas provided in Tables B1 and B2, illustrated for specific parameter values ($\theta = 0.6$, $\alpha = 0$, $\mu = 1$, $c = 0.1$, $H = 1$). Where the dotted and solid curves overlap, the solid curve represents both industry structures. See Figure 1 for additional notes.

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

Proof. The proof for variables P_c^*, Π_c^*, and MSB_c^*, which are constant over relevant intervals of $\mathcal{R}_0$, are obvious from Table B1.

The reader can verify that Q_c^* is continuous at the boundaries between cases in Table B1. The table entries for Q_c^* are constant for all cases except (LR3). In this non-empty case, Q_c^* is strictly increasing in $\mathcal{R}_0$. Variables MPB_c^* and W_c^* are analyzed similarly.

The reader can verify that $I_\infty(Q_c^*)$ is continuous at the boundary between cases in the table and further verify that $I_\infty(Q_c^*)$ is constant in $\mathcal{R}_0$ in case (LR1), increasing in case (LR2), decreasing in case (LR3), and increasing in cases (LR4) and (LR5). A local optimum is thus attained at the boundary between cases (LR2) and (LR3). Substituting $\mathcal{R}_0 = 1/(1-\tilde{c})$ in the case (LR2) table entry yields $I_\infty(Q_c^*) = \tilde{c}\mu/(\alpha+\mu)$. The only other candidate for a global maximum is

$$\lim_{\mathcal{R}_0 \uparrow \infty} I_\infty(Q_c^*) = \frac{(1-\theta)\mu}{\alpha+\mu}. \tag{B55}$$

This is less than the local optimum on the boundary between cases (LR2) and (LR3) if and only if $\tilde{c} > 1-\theta$.

The reader can verify that MEX_c^* is continous at the case (LR2)–(LR3) boundary and at the case (LR3)–(LR4) boundary and further verify that MEX_c^* is decreasing or constant in $\mathcal{R}_0$ in cases (LR2)–(LR5). Hence, MEX_c^* approaches its supremum $\lim_{\mathcal{R}_0 \downarrow 1} MEX_c^* = 1$ at the boundary between cases (LR1) and (LR2). *Q.E.D.*

Monopoly Equilibrium

Since a monopolist charges a markup above cost, $P_m^* \geq c = P_c^*$, implying $Q_m^* \leq Q_c^*$. Thus, in cases (LR1) and (LR2) in which $Q_c^* = 0$, we have $Q_m^* = 0$. Cases (LR1) and (LR2) are thus trivially identical across perfect competition and monopoly. In the remaining cases, competitive firms are able to make positive sales at price c. By continuity, the monopolist can make positive sales at some small markup above c, implying $Q_m^* > 0$ for $\mathcal{R}_0$ in case (LR3) and above.

To solve for Q_m^* in these other cases, we proceed as in the short-run analysis. The monopolist maximizes profit—which as shown in the short-run analysis can be written $[MPB(Q)-c]Q$—subject to the constraint $Q \leq \mu$ that no more than the population of newborn consumers can be served. Applying the Kuhn-Tucker method yields the following solution. If $\mathcal{R}_0 < 1/(1-\theta)^2(1-\tilde{c})$, then the constraint does not bind, yielding solution

$$Q_m^* = \frac{\mu}{\theta}\left[1 - \sqrt{\frac{1}{(1-\tilde{c})\mathcal{R}_0}}\right]. \tag{B56}$$

Otherwise, $Q_m^* = \mu$. For the constraint to bind requires the vaccine to be imperfectly effective ($\theta < 1$). The monopolist would never sell a perfectly effective vaccine to all consumers because this would eradicate the disease, leaving the monopolist with zero steady-state profit according to Proposition B2. Substituting Q_m^* into the formulas supplied for the other variables yields the entries in Table B2. The analysis is straigthforward, so we omit formal verification.

The next proposition distills the comparative-static effects of an increase in $\mathcal{R}_0$ on the steady-state equilibrium under monopoly from the entries in Table B2.

Table B2: Long-Run Equilibrium Variables under Monopoly as Functions of $\mathcal{R}_0$

	Case			
	(LR1)	(LR2)	(LR3), (LR4)	(LR5)
Variable	$\mathcal{R}_0 \in [0,1]$	$\mathcal{R}_0 \in \left(1, \frac{1}{1-\tilde{c}}\right]$	$\mathcal{R}_0 \in \left(\frac{1}{1-\tilde{c}}, \frac{1}{(1-\theta)^2(1-\tilde{c})}\right]$	$\mathcal{R}_0 \in \left(\frac{1}{(1-\theta)^2(1-\tilde{c})}, \infty\right)$
P_m^*	†	†	$\theta H\left(1-\sqrt{\frac{1-\tilde{c}}{\mathcal{R}_0}}\right)$	$\theta H\left[1-\frac{1}{(1-\theta)\mathcal{R}_0}\right]$
Q_m^*	0	0	$\frac{\mu}{\theta}\left[1-\frac{1}{\sqrt{(1-\tilde{c})\mathcal{R}_0}}\right]$	μ
Π_m^*	0	0	$\frac{\mu H(1-\tilde{c})}{\theta}\left[1-\frac{1}{\sqrt{(1-\tilde{c})\mathcal{R}_0}}\right]^2$	$\mu\theta H\left[1-\tilde{c}-\frac{1}{(1-\theta)\mathcal{R}_0}\right]$
$I_\infty(Q_m^*)$	0	$\frac{\mu}{\alpha+\mu}\left(1-\frac{1}{\mathcal{R}_0}\right)$	$\frac{\mu}{\alpha+\mu}\left[\frac{1}{\sqrt{(1-\tilde{c})\mathcal{R}_0}}-\frac{1}{\mathcal{R}_0}\right]$	$\frac{\mu}{\alpha+\mu}\left(1-\theta-\frac{1}{\mathcal{R}_0}\right)$
MPB_m^*	0	$\theta H\left(1-\frac{1}{\mathcal{R}_0}\right)$	$\theta H\left(1-\sqrt{\frac{1-\tilde{c}}{\mathcal{R}_0}}\right)$	$\theta H\left[1-\frac{1}{(1-\theta)\mathcal{R}_0}\right]$
MSB_m^*	0	θH	θH	θH
MEX_m^*	0	$\frac{\theta H}{\mathcal{R}_0}$	$\theta H\sqrt{\frac{1-\tilde{c}}{\mathcal{R}_0}}$	$\frac{\theta H}{(1-\theta)\mathcal{R}_0}$
W_m^*	$H(\alpha+\mu)$	$H\left(\alpha+\frac{\mu}{\mathcal{R}_0}\right)$	$H\left[\alpha+\mu\left(1-\tilde{c}+\frac{1}{\mathcal{R}_0}-\sqrt{\frac{1-\tilde{c}}{\mathcal{R}_0}}\right)\right]$	$H\left[\alpha+\mu\theta(1-\tilde{c})+\frac{\mu}{\mathcal{R}_0}\right]$

Notes: The distinction between cases (LR3) and (LR4) in the penultimate column, relevant for perfect competition in previous table, is irrelevant for monopoly here. †Any value $P_m^* \geq c$ is consistent with zero sales in equilibrium.

Proposition B4. *Consider the comparative-static effect of $\mathcal{R}_0$ on steady-state equilibrium under monopoly.*

- *Q_m^*, MPB_m^*, and Π_m^* are weakly increasing in $\mathcal{R}_0$ for all $\mathcal{R}_0 > 0$.*
- *A sequence of equilibrium prices exists for which P_m^* is weakly increasing in $\mathcal{R}_0$ for all $\mathcal{R}_0 > 0$.*
- *If $\tilde{c} < 1/2$ and $\tilde{c} \leq (1-2\theta)/2(1-\theta)$, then $I_\infty(Q_m^*)$ is weakly increasing in $\mathcal{R}_0$ for all $\mathcal{R}_0 > 0$. Otherwise, $I_\infty(Q_m^*)$ is nonmonotonic, reaching a local maximum at $\mathcal{R}_0 = 1/(1-\tilde{c})$ if $\tilde{c} \geq 1/2$, which is a global maximum if $\tilde{c} \geq 1-\theta$, and reaching a local maximum at $\mathcal{R}_0 = 4(1-\theta)$ if $\tilde{c} < 1/2$, which is a global maximum if $\tilde{c} \geq (3-4\theta)/4(1-\theta)$.*
- *Marginal social benefit is constant for all $\mathcal{R}_0 > 1$, with $MSB_m^* = \theta H$.*
- *The marginal externality is nonmonotonic for $\mathcal{R}_0 > 0$, approaching the supremum $MEX_m^* = \theta H$ as $\mathcal{R}_0 \downarrow 1$.*
- *If $\theta \leq 1/4$, then W_m^* is weakly decreasing in $\mathcal{R}_0$ for all $\mathcal{R}_0 > 0$. Otherwise, W_m^* is nonmonotonic in $\mathcal{R}_0$, reaching a local minimum at $\mathcal{R}_0 = 4/(1-\tilde{c})$, which is a global minimum if $\theta \geq 3/4$.*

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

Proof. The proof for MSB_m^*, which is constant within relevant intervals of $\mathcal{R}_0$, is obvious from Table B2. 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.

To derive comparative statics for P_m^*, form a price function out of the correspondence in Table B2 by taking the determinate prices in cases (LR3)–(LR5) and extending this function by taking $P_m^* = c$ in cases (LR1) and (LR2) in which the equilibrium monopoly price is indeterminate. One can then proceed to analyze the comparative statics of this price function in a similar way to variables in the preceding paragraph.

To derive comparative statics for $I_\infty(Q_m^*)$, the reader can verify that $I_\infty(Q_m^*)$ is continuous at the boundary between cases in Table B2 and further verify that $I_\infty(Q_m^*)$ is constant in $\mathcal{R}_0$ in case (LR1), increasing in case (LR2), and increasing in case (LR5). This leaves cases (LR3) and (LR4). Differentiating the table entry for those cases,

$$\frac{\partial I_\infty(Q_m^*)}{\partial \mathcal{R}_0} = \frac{\mu}{(\alpha+\mu)\mathcal{R}_0^2}\left(1-\frac{1}{2}\sqrt{\frac{\mathcal{R}_0}{1-\tilde{c}}}\right). \tag{B57}$$

Equation (B57) is decreasing in $\mathcal{R}_0$. It is negative for all $\mathcal{R}_0$ in cases (LR3) and (LR4) if it is nonpositive at the lower boundary of case (LR3). Evaluating (B57) at this lower boundary $\mathcal{R}_0 = 1/(1-\tilde{c})$, we see it is nonpositive if and only if $\tilde{c} \geq 1/2$. But then $I_\infty(Q_m^*)$ is nonmonotonic, for it is increasing in case (LR2) and decreasing in cases (LR3) and (LR4). It reaches a local maximum of $\tilde{c}\mu/(\alpha+\mu)$ at the boundary between cases (LR2) and (LR3). This weakly exceeds $\lim_{\mathcal{R}_0\uparrow\infty} I_\infty(Q_m^*) = (1-\theta)\mu/(\alpha+\mu)$ if $\tilde{c} \geq 1-\theta$, in which case the local maximum is a global maximum.

Assume $\tilde{c} < 1/2$. Then (B57) is nonnegative for all $\mathcal{R}_0$ in cases (LR3) and (LR4) if and only if (B57) is nonnegative at the upper boundary of case (LR4). Evaluating (B57) at this upper boundary $\mathcal{R}_0 = 1/(1-\theta)^2(1-\tilde{c})$, we see it is nonnegative if and only if $\tilde{c} \leq (1-2\theta)/2(1-\theta)$. Under these conditions, $I_\infty(Q_m^*)$ is nondecreasing in $\mathcal{R}_0$ for all $\mathcal{R}_0 > 0$.

For the remaining parameters, $I_\infty(Q_m^*)$ reaches a local maximum in the interior of cases (LR3) and (LR4). Setting (B57) to 0 and solving implies that the local maximum is at $\mathcal{R}_0 = 4(1-\tilde{c})$. This local maximum equals $\mu/4(\alpha+\mu)(1-\tilde{c})$, which exceeds $\lim_{\mathcal{R}_0\uparrow\infty} I_\infty(Q_m^*) = (1-\theta)\mu/(\alpha+\mu)$ if $\tilde{c} \geq (3-4\theta)/4(1-\theta)$, in which case the local maximum is a global maximum.

To derive comparative statics for W_m^*, the reader can verify that W_m^* is continuous at the boundary between cases in Table B2 and further verify that W_m^* is constant in $\mathcal{R}_0$ in case (LR1), decreasing in case (LR2), and decreasing in case (LR5). This leaves cases (LR3) and (LR4). Differentiating the table entry for those cases,

$$\frac{\partial W_m^*}{\partial \mathcal{R}_0} = \frac{\mu H}{\mathcal{R}_0^2}\left(\frac{1}{2}\sqrt{(1-\tilde{c})\mathcal{R}_0}-1\right). \tag{B58}$$

The sign of (B58) depends on the factor in parentheses, which is increasing in $\mathcal{R}_0$. It is nonpositive for all $\mathcal{R}_0$ in cases (LR3) and (LR4) if and only if it is nonpositive at the upper boundary of case (LR4). Evaluating (B57) at this upper boundary $\mathcal{R}_0 = 1/(1-\theta)^2(1-\tilde{c})$, we see it is nonpositive if and only if $\theta \leq 1/4$. Otherwise, W_m^* reaches a local minimum in the interior of cases (LR3) and (LR4).

Setting (B58) to 0 and solving yields a local minimum at $\mathcal{R}_0 = 4/(1-\tilde{c})$. This local minimum equals $H[\alpha+(3/4)(1-\tilde{c})\mu]$, which weakly exceeds $\lim_{\mathcal{R}_0\uparrow\infty} W_m^* = H[\alpha+\mu\theta(1-\tilde{c})]$ if $\theta \geq 3/4$, in which case the local minimum is a global minimum. *Q.E.D.*

To help parse the complicated conditions behind the nonmonotonicity of $I_\infty(Q_m^*)$ and W_m^* in the proposition, note that if $\theta > 3/4$, then $I_\infty(Q_m^*)$ has an interior global maximum and W_m^* has an interior global minimum in $\mathcal{R}_0$. In practice, agencies such as the U.S. Food and Drug Administration (FDA) do not typically approve vaccines with efficacy below 80% (Brennan 2009). Therefore, for realistic parameter values we have the surprising result that if a vaccine is sold by a monopolist directly to consumers, disease prevalence will be greatest and social welfare lowest not for diseases with the most extreme infectiveness as indexed by $\mathcal{R}_0$ but for moderate infectiveness.

Government Subsidies

Before analyzing particular market structures, we provide some general principles behind the optimal subsidy GS^{**} that apply to any market structure, presented in the next proposition.

Proposition B5. *Consider any market structure in which a subsidy increase weakly reduces the equilibrium price induced. If $\mathcal{R}_0 \leq 1$, then $GS^{**} = 0$; the disease is eradicated in the steady state without a vaccine or subsidy. If $\mathcal{R}_0 \in (1, 1/(1-\theta))$, then GS^{**} induces equilibrium quantity $Q^{**} = \mathcal{Q}_0$ and price $P^{**} = 0$, resulting in eradication of the disease in the steady state. If $\mathcal{R}_0 > 1/(1-\theta)$, then GS^{**} induces equilibrium quantity $Q^{**} = \mu$ and price*

$$P^{**} = \theta H\left[1 - \frac{1}{(1-\theta)\mathcal{R}_0}\right], \tag{B59}$$

resulting in universal vaccination but not eradication.

Proof. Suppose $\mathcal{R}_0 \leq 1$. Then $Q^{**} = \min(\mathcal{Q}_0, \mu) = \min(0, \mu) = 0$, where the first equality follows from equation (B50), the second from substituting $\mathcal{R}_0 \leq 1$ into (B39), and the third from $\mu > 0$. No subsidy is needed to generate zero quantity, implying $GS^{**} = 0$.

Suppose $1 < \mathcal{R}_0 < 1/(1-\theta)$. Substituting $1 < \mathcal{R}_0$ into (B39) yields $\mathcal{Q}_0 > 0$, implying $Q^{**} = \min(\mathcal{Q}_0, \mu) > 0$. Substituting $\mathcal{R}_0 < 1/(1-\theta)$ into (B39) yields $\mathcal{Q}_0 < \mu$, implying $Q^{**} = \min(\mathcal{Q}_0, \mu) = \mathcal{Q}_0$, in turn implying $I^{**} = 0$ by (B35). By Proposition B2, $I^{**} = 0$ and $Q^{**} > 0$ imply $P^{**} = 0$.

Suppose $\mathcal{R}_0 \geq 1/(1-\theta)$. Then $\mathcal{Q}_0 \geq \mu$ by (B39), implying $Q^{**} = \min(\mathcal{Q}_0, \mu) = \mu$. 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 $P' > MPB(\mu)$. Then $D(P') \leq d(P')$ since only the first two branches of (B43) are relevant in the computation of $D(P')$. Combining the inequality

$$P' > MPB(\mu) = \frac{\theta H}{1-\theta}\left(1 - \theta - \frac{1}{\mathcal{R}_0}\right) \tag{B60}$$

with (B44) 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 $MPB(\mu)$, we have $P^{**} \leq MPB(\mu)$. The highest price satisfying $P^{**} \leq MPB(\mu)$ is $P^{**} = MPB(\mu)$. To verify that this price yields the desired quantity, we have $\mu = d(MPB(\mu)) = D(MPB(\mu)) = D(P^{**})$, where the first equality follows from substituting

from (B60) into (B44) and the second equality from the fact that only the middle branch is relevant in (B43) at a price of $MPB(\mu)$.

The previous paragraph shows $d(P^{**}) = d(MPB(\mu)) = \mu$, implying $P^{**} = d^{-1}(\mu)$. Inverting (B44) yields equation (B59). *Q.E.D.*

To gain some intuition for Proposition B5, the government would like to eradicate the disease in all circumstances if this were possible. When $\mathcal{R}_0 \geq 1/(1-\theta)$, however, the disease is so infective 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 (B59) characterizes the highest price at which all consumers are still willing to purchase, which is associated with 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 turn to computing specific values of the optimal subsidy under the two market structures. Under perfect competition, the firms pass the subsidy directly to consumers, implying $P^{**} = c - GS_c^{**}$ and thus $GS_c^{**} = c - P^{**}$. Substituting the relevant values of P^{**} from Proposition B5 yields $GS_c^{**} = 0$ if $\mathcal{R}_0 \leq 1$, $GS_c^{**} = c$ if $\mathcal{R}_0 \in (1, 1/(1-\theta))$, and

$$GS_c^{**} = c - \theta H\left[1 - \frac{1}{(1-\theta)\mathcal{R}_0}\right] \tag{B61}$$

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

For an imperfectly effective vaccine, (B61) becomes negative for sufficiently large $\mathcal{R}_0$. Had we not ruled out negative subsidies by assumption, these negative values of (B61) would indeed constitute GS_c^{**}. Given the nonnegativity constraint on subsidies, we have $GS_c^{**} = 0$ for $\mathcal{R}_0 \geq 1/(1-\theta)(1-\tilde{c})$. 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 $\mathcal{R}_0$ 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 B5, $GS_m^{**} = 0$ if $\mathcal{R}_0 \leq 1$. If $\mathcal{R}_0 \in (1, 1/(1-\theta))$, Proposition B5 implies $Q^{**} = \mathcal{Q}_0$. Setting the quantity in (B56) derived from the monopoly's first-order condition equal to $\mathcal{Q}_0$, substituting the effective marginal cost $c - GS$ under a subsidy for c in the formula, and rearranging yields $GS_m^{**} = c + \theta H(\mathcal{R}_0 - 1)$. If $\mathcal{R}_0 \geq 1/(1-\theta)$, calculations that are similar except that the monopoly quantity in (B56) needs to equated with the relevant first-best quantity in this case, $Q^{**} = \mu$, yielding

$$GS_m^{**} = \max\left\{0, c + \theta H\left[\frac{1}{(1-\theta)^2\mathcal{R}_0} - 1\right]\right\}, \tag{B62}$$

where the max operator has been added to reflect the nonnegativity constraint on subsidies. The nonnegativity constraint binds in case (LR5)—not coincidentally the case in which the first-best would be obtained in monopoly equilibrium without a subsidy.

For reference, Table B3 reports the results just derived for the optimal subsidy under the two market structures. Figure B2 provides an illustrative graph. The next proposition catalogs relevant observations concerning GS^{**}.

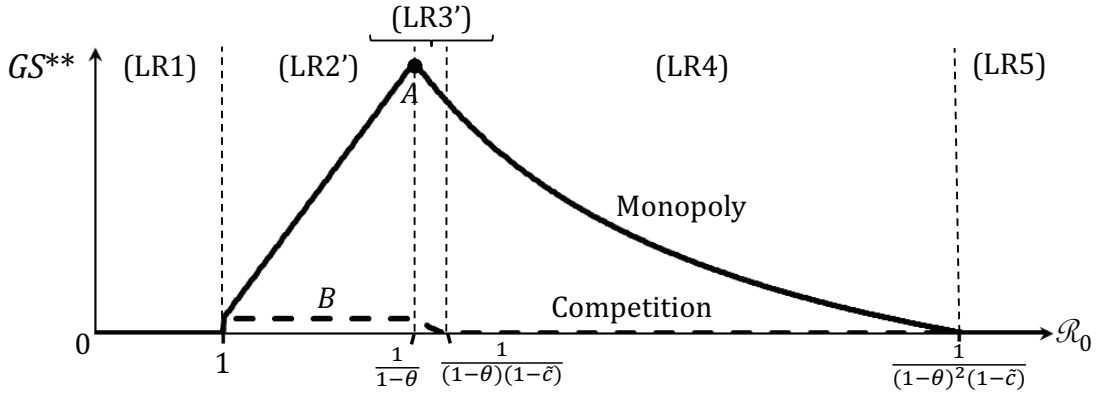
Proposition B6. *The following results characterize optimal subsidies, attaining the first best at minimum government expenditure.*

Table B3: Long-Run Optimal Subsidy GS^{**} as a Function of $\mathcal{R}_0$

	Case				
	(LR1)	(LR2′)	(LR3′)	(LR4)	(LR5)
Variable	$\mathcal{R}_0 \in [0,1]$	$\mathcal{R}_0 \in \left(1, \frac{1}{1-\theta}\right]$	$\mathcal{R}_0 \in \left(\frac{1}{1-\theta}, \frac{1}{(1-\theta)(1-\tilde{c})}\right]$	$\mathcal{R}_0 \in \left(\frac{1}{(1-\theta)(1-\tilde{c})}, \frac{1}{(1-\theta)^2(1-\tilde{c})}\right]$	$\mathcal{R}_0 \in \left(\frac{1}{(1-\theta)^2(1-\tilde{c})}, \infty\right)$
GS_c^{**}	0	c	$c+\theta H\left[\frac{1}{(1-\theta)\mathcal{R}_0}-1\right]$	0	0
GS_m^{**}	0	$c+\theta H(\mathcal{R}_0-1)$	$c+\theta H\left[\frac{1}{(1-\theta)^2\mathcal{R}_0}-1\right]$	$c+\theta H\left[\frac{1}{(1-\theta)^2\mathcal{R}_0}-1\right]$	0

Notes: Cases (LR1), (LR4), and (LR5) are identical to those in Tables B1 and B2. Cases (LR2′) and (LR3′) differ from (LR2) and (LR3) in the previous tables in that the boundary between (LR2′) and (LR3′) is $1/(1-\theta)$ rather than $1/(1-\tilde{c})$.

Figure B2: Graphs of Long-Run Optimal Subsidy GS^{**} as Function of $\mathcal{R}_0$



Notes: Graph of formulas provided in Table B3, illustrated for the specific parameter values indicated in previous figure.

- $GS_m^{**} \geq GS_c^{**}$, *with strict inequality for all* $\mathcal{R}_0 \in (1, 1/(1-\theta)^2(1-\tilde{c}))$.
- GS_c^{**} *and* GS_m^{**} *are weakly increasing in* c *and* θ.
- GS_c^{**} *and* GS_m^{**} *are nonmonotonic in* $\mathcal{R}_0 > 0$. GS_c^{**} *reaches its global maximum,* c*, for all* $\mathcal{R}_0$ *in the interval* $(1, 1/(1-\theta)]$*;* GS_m^{**} *reaches its global maximum,* $c+H\theta^2/(1-\theta)$*, for* $\mathcal{R}_0 = 1/(1-\theta)$.

Proof. We will prove the second bullet point regarding comparative statics of GS^{**} in c and θ. Gleaning the remaining results from Table B3 is relatively straightforward. In case (LR1) of the table, $GS_c^{**} = GS_m^{**} = 0$, in which case the comparative statics hold trivially. The remaining cases can be combined in a single expression for each market structure,

$$GS_c^{**} = \min\left\{c, \max\left\{0, c+\theta H\left[\frac{1}{(1-\theta)\mathcal{R}_0} - 1\right]\right\}\right\} \tag{B63}$$

$$GS_m^{**} = \min\left\{c+\mathcal{R}_0-1, \max\left\{0, c+\theta H\left[\frac{1}{(1-\theta)^2\mathcal{R}_0} - 1\right]\right\}\right\}, \tag{B64}$$

which are obviously weakly increasing in c and θ. *Q.E.D.*

Limiting Results

While Tables B1 and B2 help organize the results, their numerous entries and complex formulas can be further distilled by taking limits of certain parameters justified by pedagogical or practical grounds. 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. For example, the 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. For rescaled cost, $\tilde{c} = c/\theta H$, following the logic of Section 7, we take the limiting case of a costless vaccine, $\tilde{c} \downarrow 0$. 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 $\mathcal{R}_0 > 1$ considerably simplifies the analysis. Only case (LR3) remains from Tables B1 and B2 and case (LR2′) from Table B3. The following proposition is then immediate from inspection of the tables.

Proposition B7. *Suppose $\mathcal{R}_0 > 1$ and consider the limits $\theta \uparrow 1$, $\tilde{c} \downarrow 0$, and $\alpha \downarrow 0$.*

- *Under perfect competition, $P_c^* = MPB_c^* = \Pi_c^* = I_\infty(Q_c^*) = GS_c^{**} = 0$, $Q_c^* = \mu(1-1/\mathcal{R}_0)$, and $MEX_c^* = H$.*
- *Under monopoly, $P_m^* = MPB_m^* = H(1-1/\sqrt{\mathcal{R}_0})$, $Q_m^* = \mu(1-1/\sqrt{\mathcal{R}_0})$, $\Pi_m^* = \mu H(1-1/\sqrt{\mathcal{R}_0})^2$, $I_\infty(Q_m^*) = (1/\sqrt{\mathcal{R}_0}) - (1/\mathcal{R}_0)$, $MEX_m^* = H/\sqrt{\mathcal{R}_0}$, and $GS_m^{**} = H(\mathcal{R}_0 - 1)$.*

According to the proposition, under perfect competition, for all $\mathcal{R}_0 > 1$, 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 $\mathcal{R}_0 \downarrow 1$, while price approaches 100% of the harm from contracting the disease and quantity approaches 100% share of newborns in the limit $\mathcal{R}_0 \uparrow \infty$. Equilibrium prevalence is nonmonotonic in $\mathcal{R}_0$ under monopoly. Maximizing the formula given for $I_\infty(Q_m^*)$, one can show that equilibrium prevalence is greatest for a disease with $\mathcal{R}_0 = 4$.

Calibrations

This section provides calibrations for a series of relevant diseases to illustrate the implications of the long-run analysis. As noted in the ealier calibration exercise for Covid, the calibrations are meant more as illustrations than forecasts, and the caveats issued there continue to apply. The calibrations use the limiting parameter values assumed in the previous section, allowing us to apply those limiting results.

HIV Calibration: Our first calibration considers a disease, HIV, with a moderate $\mathcal{R}_0$. Anderson and May (1991) cite estimates of $\mathcal{R}_0$ ranging from 2 to 5 for Type-1 HIV among men who have sex with men (MSM). For convenience, we take the round number $\mathcal{R}_0 = 4$ from this range. Incidentally, recall that this is the value of $\mathcal{R}_0$ for which disease prevalence is greatest in equilibrium under monopoly.

Substituting $\mathcal{R}_0 = 4$ into the formulas provided by Proposition B7, the model suggest that a monopolist 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 1/4. This is much lower than the 3/4 prevalence rate that would emerge in the absence of a vaccine for a disease with $\mathcal{R}_0 = 4$. The monopolist captures one third of the potential social surplus from a vaccine. Consumers capture one third, and the remaining third is lost to monopoly distortion.

As noted in the previous subsection, for all $\mathcal{R}_0 > 1$, including $\mathcal{R}_0 = 4$ 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.

Unfortunately, it may be unrealistic to suppose that competition would emerge even in the long run for high-tech product like a potential HIV vaccine that are extremely difficult for a generic competitor to reverse engineer. The minimum subsidy that would have to be paid to the monopolist to attain the first best is $GS_m^{**} = 3H$. In other words, the monopolist would have to receive a per-course subsidy of at least three times the lifetime harm experienced from certainly contracting HIV, easily running into 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 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.

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 need to be relaxed.

Consider the introduction of a small program subsidizing costly circumcision in a country with little current uptake. Since male circumcision is only relevant for half the population, analyzing a program targeting universal adoption would require modifying the model to allow for gendered subpopulations. We avoid this complication by supposing the program is small. Thus, instead of computing the optimal subsidy GS_c^{**} achieving universal adoption, this calibration analyzes a subsidy provided to the initial adopter.

The relevant theoretical case is (LR2), the only case with a positive infection rate but negligible adoption. The subsidy required to induce the initial participants in the small program to become circumcized in case (LR2) equals $c - MPB_c^* = c - \theta H(1 - 1/\mathcal{R}_0)$. A planner would be willing to provide such a subsidy up to the level of $MEX_c^* = \theta H/\mathcal{R}_0$. Substituting $\mathcal{R}_0 = 4$ assumed in this subsection for HIV, $\theta = 0.6$, the estimate of efficacy in randomized controlled trials in several studies including Bailey *et al.* (2007), and $H = 7,000$, the estimated lifetime cost of first-line drug treatments for HIV (UNAIDS/WHO/SACEMA Expert Group on Modelling the Impact and Cost of Male Circumcision for HIV Prevention 2009), a lower bound on the health benefit from avoiding an HIV infection, yields $MEX_c^* = 1,050$, implying that the required subsidy is somewhere between \$0 and \$1,050. Of the eight studies surveyed by Kennedy *et al.* (2020), none paid a subsidy beyond the cost of the procedure of more than \$15. While most of these studies obtained statistically significant estimates for the effect of subsidies on circumcision, Kennedy *et al.* (2020, p. 11) note that "the overall uptake ... in these studies was low, and the absolute differences between groups were small," suggesting that the subsidies were far below the optimum.

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 $\mathcal{R}_0$ than HIV. We will take $\mathcal{R}_0 = 16$ for this disease, at the low end of estimates provided by Anderson and May (1991). The model suggests that a monopolist selling a measles vaccine to consumers would price it at fully 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%, less than the 25% prevalence calibrated for HIV—as expected given the prevalence rate is maximized at the value of $\mathcal{R}_0$ used in the HIV calibration.

The minimum subsidy required to attain the first best under monopoly is $GS_m^{**} = 15H$. In other

words, the monopolist 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 $\mathcal{R}_0$ toward the higher end.

Cournot Competition

Cases (LR1) and (LR2) from Table B1, 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 (LR1) and (LR2) from both Tables B1 and B2 will also apply to Cournot.

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

$$Q_n^* = \frac{\mu}{\theta}\left[1-\sqrt{\frac{\psi}{(1-\tilde{c})\mathcal{R}_0}}\right], \tag{B65}$$

where

$$\psi = \frac{(n-1)^2}{4n^2(1-\tilde{c})\mathcal{R}_0}\left[1+\sqrt{1+\frac{2n(1-\tilde{c})\mathcal{R}_0}{(n-1)^2}}\right]^2. \tag{B66}$$

Substituting Q_n^* for Q in the relevant equations yields

$$P_n^* = MPB_n^* = \theta H\left(1-\sqrt{\frac{1-\tilde{c}}{\psi\mathcal{R}_0}}\right) \tag{B67}$$

$$I_\infty(Q_n^*) = \frac{\mu}{\alpha+\mu}\left[\sqrt{\frac{\psi}{(1-\tilde{c})\mathcal{R}_0}}-\frac{1}{\mathcal{R}_0}\right] \tag{B68}$$

$$\Pi_n^* = \frac{\mu H(1-\tilde{c})}{\theta}\left[1-\sqrt{\frac{1}{\psi(1-\tilde{c})\mathcal{R}_0}}\right]\left[1-\sqrt{\frac{\psi}{(1-\tilde{c})\mathcal{R}_0}}\right] \tag{B69}$$

$$MEX_n^* = \mu H\sqrt{\frac{1-\tilde{c}}{\psi\mathcal{R}_0}} \tag{B70}$$

$$W_n^* = H\left[\alpha+\mu\left(1-\tilde{c}+\frac{1}{\mathcal{R}_0}-\sqrt{\frac{1-\tilde{c}}{\psi\mathcal{R}_0}}\right)\right]. \tag{B71}$$

The preceding analysis is valid if

$$\mathcal{R}_0 \leq \frac{\psi}{(1-\theta)^2(1-\tilde{c})}. \tag{B72}$$

Otherwise, $Q_n^* > \mu$ for the Q_n^* in (B65). 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 variables have the same formula as in case (LR5) of Table B2. Note that the threshold between cases (LR4) and (LR5) is different, given by the right-hand side of (B72).

It is easily seen that $\psi = 1$ for $n = 1$, and thus that the preceding expressions for the equilibrium variables collapse to their monopoly values given in Table B2. It is also easily seen that $\lim_{n\uparrow\infty} \psi = 1/(1-\tilde{c})\mathcal{R}_0$, and thus that the preceding expressions for the equilibrium variables collapse to their values under perfect competition given in Table B1.

Vaccines Versus Drugs

Start by computing monopoly profit and welfare from a drug, respectively Π_{md}^* and W_{md}^*, in steady-state equilibrium in the long-run analysis. If $\mathcal{R}_0 \leq 1$, the disease naturally dies out in the steady-state, implying $Q_{md}^* = \Pi_{md}^* = 0$. If $\mathcal{R}_0 > 1$, the monopolist can charge $P_{md}^* = \theta H$ for the drug to all newly infected consumers each instant. According to equation (B30), new infections, $\beta I_\infty S_\infty$, and removals from the infected population, $(\alpha+\mu)I_\infty$, must balance each instant to maintain $\dot{I}_\infty = 0$ in the steady state. Hence, we can compute new infections as $(\alpha+\mu)I_\infty(0)$, where the argument added to $I_\infty(0)$ indicates that the drug does nothing to curtail infections, as in the vaccine model with no vaccine sales. Equilibrium drug quantity is thus $Q_{md}^* = (\alpha+\mu)I_\infty(0) = \mu(1-1/\mathcal{R}_0)$ by (B35). Assuming as in Section **??** that production is costless,

$$\Pi_{md}^* = P_{md}^* Q_{md}^* = \mu\theta H\left(1-\frac{1}{\mathcal{R}_0}\right). \tag{B73}$$

Welfare is

$$W_{md}^* = h[1-I_\infty(0)+\theta I_\infty(0)] = H\left[\alpha+\mu-\mu(1-\theta)\left(1-\frac{1}{\mathcal{R}_0}\right)\right]. \tag{B74}$$

Comparing these expressions against the analogous entries in Table B2 for a vaccine leads to the next proposition.

Proposition B8. *Consider the long-run analysis with $c = 0$. For all $\mathcal{R}_0 > 0$, $\Delta W^{**} \leq 0$ and $\Delta\Pi_m^* \geq 0$, with strict inequality if and only if $\mathcal{R}_0 > 1$. $\Delta\Pi_m^*$ is quasiconcave in $\mathcal{R}_0$, with $\lim_{\mathcal{R}_0\downarrow 0}\Delta\Pi_m^* = \lim_{\mathcal{R}_0\uparrow\infty}\Delta\Pi_m^* = \inf_{\mathcal{R}_0>0} = 0$, reaching an maximum of $\theta^2 H/(1-\theta)$ at $\mathcal{R}_0 = (1+\theta)^2$. $\Delta W_m^* < 0$ if and only if $\mathcal{R}_0 > [\theta/(1-\theta)]^2$.*

Proof. Suppose $\mathcal{R}_0 \leq 1$. 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 $\mathcal{R}_0 < 1$ implies $\lim_{\mathcal{R}_0\downarrow 0}\Delta\Pi_m^* = 0$.

Suppose $\mathcal{R}_0 > 1$. The assumption $c = 0$ implies $\tilde{c} = 0$, leaving two cases in Table B1: (LR3)–(LR4) and (LR5). In case (LR3)–(LR4), defined by

$$\mathcal{R}_0 \in \left(1, \frac{1}{(1-\theta)^2}\right), \tag{B75}$$

substituting from (B73) for Π_{md}^* and from the table entry for Π_{mv}^* yields, after rearranging,

$$\Delta\Pi_m^* = \frac{\sqrt{\mathcal{R}_0}-1}{\theta\mathcal{R}_0}\left[1+\theta-(1-\theta)\sqrt{\mathcal{R}_0}\right]. \tag{B76}$$

The first factor is positive since $\sqrt{\mathcal{R}_0} > 1$ by (B75); the second is positive since $\sqrt{\mathcal{R}_0}(1-\theta) < 1$ by (B75). In case (LR5), $\Delta\Pi_m^* = \theta/(1-\theta)\mathcal{R}_0$, implying $\lim_{\mathcal{R}_0\uparrow\infty} \Delta\Pi_m^* = 0$. Combined with the results for $\mathcal{R}_0 \leq 1$, we have that $\Delta\Pi_m^* \geq 0$ for all $\mathcal{R}_0 > 0$ with strict inequality if and only if $\mathcal{R}_0 > 1$. Further, $\inf_{\mathcal{R}_0>0} = 0$.

To verify the quasiconcavity of $\Delta\Pi_m^*$, $\Delta\Pi_m^*$ is a constant 0 in (LR1). In case (LR3)–(LR4),

$$\frac{\partial \Delta\Pi_m^*}{\partial \mathcal{R}_0} = \frac{H}{\mathcal{R}_0^2}\left(1+\theta-\sqrt{\mathcal{R}_0}\right), \tag{B77}$$

implying that $\Delta\Pi_m^*$ is first increasing, reaches a critical point at $\mathcal{R}_0 = (1+\theta)^2$, and then is decreasing. One can verify that the critical point is in the interior of case (LR3)–(LR4), as $(1+\theta)^2 < 1/(1-\theta)^2$. At the boundary of (LR5), $\Delta\Pi_m^*$ is continuous and continues to decline throughout case (LR5), proving $\Delta\Pi_m^*$ is quasiconcave for all $\mathcal{R}_0 > 0$.

Turning to ΔW_m^*, in case (LR5), substituting from (B74) for W_{md}^* and from the relevant table entry for W_{mv}^* yields $\Delta W_m^* = -\theta h/\mathcal{R}_0$, which is negative. In case (LR3)–(LR4), substituting the relevant table entry for W_{mv}^* in case (LR3)–(LR4) yields

$$\Delta W_m^* = \frac{h}{\mathcal{R}_0}\left[(\theta-1)(\sqrt{\mathcal{R}_0})^2+\sqrt{\mathcal{R}_0}-\theta\right]. \tag{B78}$$

The sign is determined by the factor in brackets, a quadratic equation in $\sqrt{\mathcal{R}_0}$, which is negative if $\sqrt{\mathcal{R}_0}$ lies outside the roots 1 and $\theta/(1-\theta)$. Since $\mathcal{R}_0 > 1$, the relevant condition is $\sqrt{\mathcal{R}_0} > \theta/(1-\theta)$, implying $\mathcal{R}_0 > [\theta/(1-\theta)]^2$, 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 (B50), the first-best vaccine quantity is $Q^{**} = \mathcal{Q}_0 = (\mu/\theta)(1-1/\mathcal{R}_0)$ if $\mathcal{R}_0 \in (1, 1/(1-\theta)]$ and $Q^{**} = \mu$ if $\mathcal{R}_0 > 1/(1-\theta)$. Suppose $\mathcal{R}_0 \in (1, 1/(1-\theta)]$. Then $W_v^{**} = (\alpha+\mu)H$. Substituting this value along with the value of $W_d^* = W_d^{**}$ from (B74) and rearranging yields $\Delta W^{**} = W_d^{**} - W_v^{**} = -\mu H(1-\theta)(1-1/\mathcal{R}_0) < 0$. Next, suppose $\mathcal{R}_0 > 1/(1-\theta)$. Then, according to the case (LR5) entry in Table B2 setting $\tilde{c} = 0$, $W_v^{**} = H(\alpha+\theta\mu+\mu/\mathcal{R}_0)$. Substituting this value along with the value of $W_d^* = W_d^{**}$ from (B74) and rearranging yields $\Delta W^{**} = W_d^{**} - W_v^{**} = -\theta\mu H/\mathcal{R}_0$. Thus, $\Delta W^{**} \leq 0$ for all $\mathcal{R}_0 > 0$ with strict inequality for $\mathcal{R}_0 > 1$. *Q.E.D.*

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 ($\mathcal{R}_0 > 1$). However, if $\mathcal{R}_0 > [\theta/(1-\theta)]^2$, social welfare is higher with a vaccine, implying that the monopoly is biased toward the "wrong" product for sufficiently high $\mathcal{R}_0$. While the drug has the advantage that the monopolist 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 $\mathcal{R}_0$, greatest for some interior value. The externality disappears if the disease is noninfective and is swallowed by consumers' private benefit if the disease is infinitely infective.

Consumer Heterogeneity

The analysis of consumer heterogeneity in Appendix B3 for the short-run model through Lemma B1 can be repeated for the long-run model with the exception that where the trailing factor $\hat{S}_0$ appears in

equations (B25)–(B27), μ needs to be substituted. It remains to complete the proof of Proposition B1 in the long-run analysis by showing that the first factor in equation (B27) approaches 0 in both limits $\mathcal{R}_0 \downarrow 0$ and $\mathcal{R}_0 \uparrow \infty$.

One can use direct calculation to verify that the formula for $\Phi_I(Q)$ derived for homogeneous consumers in (B42) continues to hold with consumers heterogeneous in harm. For $\mathcal{R}_0 \leq 1$, $\mathcal{Q}_0 = 0$, implying $\Phi_I(Q) = 0$ for all Q, implying $\partial\Phi_I(Q)/\partial Q = 0$ for all Q, implying $\lim_{\mathcal{R}_0 \downarrow 0} \partial\Phi_I(Q^*)/\partial Q = 0$ for both $Q^* = Q_c^*$ and $Q^* = Q_m^*$.

For $\mathcal{R}_0 \uparrow \infty$, $\mathcal{Q}_0 = \mu/\theta > \mu$ for $\theta < 1$, implying the second branch of (B42) is the relevant one. Differentiating,

$$\frac{\partial\Phi_I(Q)}{\partial Q} = \frac{\theta}{\mu\mathcal{R}_0(1-\theta Q/\mu)^2}. \tag{B79}$$

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

$$\frac{\theta}{\mu\mathcal{R}_0} \leq \frac{\partial\Phi_I(Q)}{\partial Q} \leq \frac{\theta}{\mu\mathcal{R}_0(1-\theta)^2}. \tag{B80}$$

Taking limits,

$$\lim_{\mathcal{R}_0\uparrow\infty} \frac{\theta}{\mu\mathcal{R}_0} \leq \lim_{\mathcal{R}_0\uparrow\infty} \frac{\partial\Phi_I(Q)}{\partial Q} \leq \lim_{\mathcal{R}_0\uparrow\infty} \frac{\theta}{\mu\mathcal{R}_0(1-\theta)^2}. \tag{B81}$$

The limits on the far left and far right-hand sides of (B81) equal 0, implying $\lim_{\mathcal{R}_0\uparrow\infty} \partial\Phi_I(Q)/\partial Q = 0$ for all $Q \in [0, \mu]$, including $Q = Q_c^*$ and $Q = Q_m^*$.

Additional References

Bailey, Robert C., Stephen Moses, Corette B. Parker, Kawango Agot, Kian Maclean, John N. Krieger, Carolyn F. M. Williams, Richard T. Campbell, and Jeckoniah O. Ndinya-Achola. (2007) "Male Circumcision for HIV Prevention in Young Men in Kisumu, Kenya: A Randomised Controlled Trial," *Lancet* 369: 643–656.

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

Kennedy, Caitlin E., Ping T. Yeh, Kaitlyn Atkins, Virginia A. Fonner, Michael D. Sweat, Kevin R. O'Reilly, George W. Rutherford, Rachel Baggaley, and Jluia Samuelson. (2020) "Economic Compensation Interventions to Increase Uptake of Voluntary Medical Male Circumcision for HIV Prevention: A Systematic Review and Meta-Analysis," *PLoS ONE* 15(1): e0227623.

UNAIDS/WHO/SACEMA Expert Group on Modelling the Impact and Cost of Male Circumcision for HIV Prevention. (2009) "Male Circumcision for HIV Prevention in High HIV Prevalence Settings: What Can Mathematical Modelling Contribute to Informed Decision Making?" *PLoS Medicine* 6(9): e1000109.

U.S. Centers for Disease Control and Prevention. (2020) "Vaccines and Preventable Diseases: Vaccines by Disease." [Internet.] Vaccine efficacies downloaded August 28, 2020 from https://www.cdc.gov/vaccines/pubs/pinkbook/hepb.html for hepatitis B, from https://www.cdc.gov/vaccines/vpd/mmr/public/index.html for measles, from https://www.cdc.gov/vaccines/vpd/dtap-tdap-td/hcp/about-vaccine.html for pertussis and tetanus, from https:

//www.cdc.gov/vaccines/vpd/polio/index.html for polio, and from from https://www.cdc.gov/vaccines/vpd/shingles/index.html for shingles.

Ward, Courtney J. (2014) "Influenza Vaccination Campaigns: Is an Ounce of Prevention Worth a Pound of Cure?" *American Economic Journal: Applied Economics* 6: 38–72.