OPTIMAL VACCINE SUBSIDIES FOR EPIDEMIC DISEASES

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

Abstract—We analyze optimal vaccine subsidies in a model integrating disease epidemiology into a market with rational economic agents. The focus is on an intensive vaccine campaign to quell an epidemic in the short run. Across a range of market structures, positive vaccine externalities and optimal subsidies peak for diseases that spread quickly, but not so quickly that everyone is driven to be vaccinated. We assess the practical relevance of this peak—as well as the existence of increasing social returns to vaccination and optimality of universal vaccination—in calibrations to the COVID-19 pandemic.

I. Introduction

Technologies such as vaccines, condoms, and mosquito nets that protect individuals against infectious diseases can generate positive externalities by reducing transmission to others. While textbook economic models provide a justification for public subsidies of such preventive technologies, the appropriate magnitude of these subsidies remains understudied. Without a deeper understanding, it is difficult for economists to provide guidance—even at a conceptual level—on the optimal level of government subsidies for infectious-disease control and how such subsidies should vary across diseases. Furthermore, much existing work by economists has been oriented toward policies combating endemic diseases playing out over generations, such as vaccination campaigns to eradicate polio or circumcision campaigns to reduce the spread of HIV. The COVID-19 pandemic, however, has underscored the urgency of quelling global outbreaks of novel diseases.

To address these questions, we construct a tractable model that integrates epidemiological and economic considerations. For concreteness, we focus on the market for a vaccine, but the analysis applies to other 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. We analyze a vaccine campaign introduced at a single point in time into a population without turnover, modeling choices intended to capture relevant features of epidemics like COVID-19 that may rise and fall within a generation, calling for a concentrated policy response.

A key finding is that the positive externality that the marginal consumer’s vaccination exerts on others reaches its peak for intermediate rather than the highest values of $R_0$ (the disease’s basic reproductive number, a widely used measure of infectiousness). This finding holds robustly across market structures ranging from perfect competition to Cournot competition to monopoly. For low values of $R_0$, the marginal externality is low because there is little disease transmission among people. For high values of $R_0$, vaccinating a given consumer does not provide much protection to others since they are almost certain to contract the disease from someone else anyway. To be sure, a consumer’s vaccination provides a substantial social benefit when $R_0$ is extremely high, but most of that benefit is internalized by the consumer. The optimal subsidy, which corrects for the marginal externality, likewise peaks for intermediate values of $R_0$.

At certain intermediate values of $R_0$, free riding can be so extensive that an increase in $R_0$, by reducing some of this free riding, can perversely reduce equilibrium infections. This epidemiological version of the Peltzmann (1975) effect arises when consumers compensate for increased risk by increasing vaccination to such an extent that it more than offsets the direct increase in infectiousness.

Our results go beyond nonmonotonicities. Whether universal vaccination can be a viable business strategy is explored in section V. Previous game-theoretic analyses have suggested 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 & Phillipson, 1997; May, 2000; Bauch & Earn, 2004). In our model, however, universal vaccination of susceptibles 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 possible but guaranteed in equilibrium for sufficiently low cost and sufficiently high infectiousness.

In typical economic settings, benefit functions are concave, leading to decreasing returns, but epidemiological effects may lead to disproportionate benefits of vaccination, as folk wisdom suggests may happen around the threshold for herd immunity. Section VI explores this issue formally by providing conditions under which vaccination exhibits increasing social returns. The key indicator is \( R_0 \hat{S}_0 \), the product of the basic reproductive number \( R_0 \) (technically, the number of secondary cases an infectious individual would hypothetically transmit in a fully susceptible population) and the proportion of susceptibles in the relevant population \( \hat{S}_0 \) (here, at date 0, when the vaccine is introduced). The product \( R_0 \hat{S}_0 \), known as the effective reproductive number, tells us the number of secondary cases an infectious individual transmits in the relevant population. If \( R_0 \hat{S}_0 < 1 \), primary infections lead to fewer secondary infections from the start, implying the infection rate falls throughout the epidemic even absent a vaccine. For infections to rise to a peak before falling in the classic epidemic pattern requires \( R_0 \hat{S}_0 > 1 \). If the effective reproductive number is yet higher—we show \( R_0 \hat{S}_0 > 2 \) is sufficient—then social returns to vaccination are initially increasing. With such a high effective reproductive number, the epidemic is so explosive that a small amount of vaccine does little to slow it; to make a measurable dent in the next.

Section VII compares the results to a market for a drug that is similar in all ways to the vaccine except that it treats symptoms but does nothing to reduce disease transmission from treated individuals. We show that a monopoly always prefers to develop the drug, but parameters exist for which social welfare is higher with the vaccine. Consistent with nonmonotonicities found elsewhere, the monopoly’s bias toward a drug peaks for intermediate values of \( R_0 \).

Section VIII calibrates the model to the COVID-19 pandemic. While too stylized for quantitative policy guidance, the calibrations provide qualitative insights into the magnitude of distortions caused by externalities and market power and allow us to assess whether the theoretical conditions behind results such as the Peltzman effect, increasing social returns, and a bias toward drugs versus vaccines are practically relevant.

Our analysis is intentionally built on a basic epidemiological model, sacrificing realism to obtain rigorous propositions involving interpretable economic conditions rather than results based on isolated simulations or structural estimates tied to current circumstances. The SIR model omits features required for quantitative forecasting in a real-world epidemic such as heterogeneous agents (Ellison, 2020; Acemoglu et al., 2021), transmission along networks (Newman, 2002; Fajgelbaum et al., 2021), and macroeconomic dynamics (Eichenbaum, Rebelo, & Trabandt, 2020). Perhaps the key omission is endogenous social distancing, which would flatten the epidemic’s path relative to SIR predictions. A growing literature has advanced increasingly sophisticated models to approximate the dynamic behavior of rational, forward-looking agents, seeking to reduce risk by curtailing activity. We do not address all of these omissions but address some in several pieces of additional work. Online appendix A.4 allows consumers to be heterogeneous in harm. Online appendix A.5 extends the model to allow consumers to purchase a second preventive in addition to, or instead of, the vaccine. While not capturing a continuously updating distancing decision, the extension could capture fixed investments in masks or lifestyle changes. Our companion paper (Goodkin-Gold et al., 2022) maintains the SIR framework but adopts alternative modeling assumptions suited to an endemic disease, incorporating population turnover and continuous vaccination of arriving newborn cohorts. The analysis focuses on the steady-state equilibrium in which the effective reproductive number is always 1, which Gans (2020) suggests is a reasonable shortcut for modeling endogenous social distancing. All of the additional work in the online appendixes and companion paper finds that marginal externalities and optimal subsidies are highest for intermediate values of \( R_0 \), increasing confidence in the robustness of the results to modeling assumptions.

Our paper contributes to the theoretical literature analyzing vaccine externalities. Economists have long observed that vaccines may provide positive externalities that could affect consumers’ and firms’ decisions. 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 heterogeneous population and consider the effect of network structures in determining optimal vaccine allocation. Manski (2021), building on a series of the author’s earlier papers, provides guidance on optimal vaccine policies (including mandates) when the extent of externality is unknown. Avery (2021) is a recent ambitious.

---

1Recent theoretical advances include Acemoglu et al. (2020); Ateskon, Kopecky, and Zha (2020, 2021); Farboodi, Jarosch, and Shimer (2020); Jones, Philippin, and Venkateswaran (2020); Keppo et al. (2021); Makris and Toxvaerd (2020); McAdams (2020); Rachel (2020); Tröger (2020); and Toxvaerd (2019, 2020).

2See Avery et al. (2020) and McAdams (2021) for recent surveys.

3See, among others, Brito, Sheshinski, and Intrilligator (1991); Chen and Toxvaerd (2014); Francis (1997); Geoffard and Philipson (1997); Gersovitz (2003); and Gersovitz and Hammer (2004, 2005). Work in behavioral epidemiology has begun to incorporate externalities at least implicitly, considering, for example, game-theoretic analyses of decisions around whether to vaccinate or to free-ride on herd immunity (Funk, Salathé, & Jansen, 2010; Manfredi & D’Onofrio, 2013).
attempt to integrate social distancing and vaccinations in a tractable model. We contribute a precise characterization of the nonmonotonicity of externalities and optimal subsidies as a function of disease infectiousness. We also contribute by formally modeling the supply side of the vaccine market, allowing firms to have market power.

Our paper is perhaps closest to two companion papers in the operations research literature: Mamani, Adida, and Dey (2012) and Adida, Dey, and Mamani (2013). They also analyze optimal subsidies for various degrees of supplier market power. Their focus is on consumers with uniformly distributed harm. While we also examine consumer heterogeneity (see online appendix A.4), our analysis focuses on homogeneous consumers, allowing us to derive more definitive expressions for equilibrium variables, which in turn afford additional insights, allowing us to analyze the comparative-static effect of increases in $R_0$. Our central result on the nonmonotonicity of optimal subsidies in $R_0$ and other comparative statics are novel in our paper. We also provide calibrations, results on increasing social returns, and a comparison between drugs and vaccines not found in their papers.4

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 & Shattock, 2012; Keeling & Ross, 2015; Nguyen & Carlson, 2016; and Enayati & Özaltin, 2020).5 This literature has the appeal of studying increasingly rich epidemiological models but with the drawback of having to simulate results 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.

Our paper also contributes to the economic literature responding to the COVID-19 pandemic. Scholars sought to apply detailed models to forecast the course of the pandemic (Atkeson, Kopecky, & Zha, 2020), to recommend lockdown and testing protocols (Alvarez, Argente, & Lippi, 2021), and to recommend policies for prioritizing scarce vaccine supplies among heterogeneous consumers (Buckner, Chowell, & Springborn, 2021). The focus of our COVID-19 calibration is different: on optimal subsidies in a decentralized market rather than optimal strategies for a central planner. Our overall goal is also different: obtaining general principles in a stylized model rather than quantitative results in a more complex model.

4Also closely related is Althouse, Bergstrom, and Bergstrom (2010), who provide a 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 affect market outcomes and optimal policy.

5See also Anderson, Laxminarayan, and Salant (2012). Using an SIS model incorporating reinfection, the study finds that a planner prefers concentrating scarce supplies each period in the location with fewer infections.
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 $\beta I_t$. The mass of susceptibles $S_t$ generates $\beta I_t S_t$ new infections. Equation (2) indicates that the susceptible population falls by the number of newly infected and newly vaccinated. Equation (3) indicates that the infected population is increased by the number of newly infected and reduced by the mass $\alpha I_t$ of previously infected consumers who recover, where $\alpha \in (0, 1)$ denotes the recovery rate. This $\alpha I_t$ 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 $S_t$. Equation (5) reflects the instantaneous nature of the vaccination campaign, with no further vaccine administered after the initial tranche at date 0.\(^6\) We assume that if the initial vaccine course is not effective for a person, a further course will not be effective for them either. Under that assumption, administering all vaccine in the first instant is both the profit-maximizing and welfare-maximizing strategy.\(^7\)

Let $Q \in [0, \tilde{Q}]$ denote the quantity of vaccine courses administered at date 0 to susceptibles, the only consumers who can possibly benefit from vaccination. 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 the efficacy of a vaccine course. Let $\tilde{S}_0$, $\tilde{I}_0$, and $\tilde{R}_0$ denote the counterfactual value of the relevant compartments at date 0 in the absence of vaccines (so by definition, $\tilde{Z}_0 = 0$). Then the initial conditions for the SIR system can be written as

\begin{align*}
  S_0 &= \tilde{S}_0 - Z_0, & (6) \\
  I_0 &= \tilde{I}_0, & (7) \\
  R_0 &= \tilde{R}_0 = 1 - \tilde{I}_0 - \tilde{S}_0, & (8) \\
  Z_0 &= \theta Q. & (9)
\end{align*}

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

In lieu of the transmission parameter $\beta$, epidemiologists often work with a related parameter $R_0$, called the basic reproductive number, equal to the expected number of secondary cases an infectious individual transmits in a fully susceptible population. In our model,

\begin{equation}
  R_0 = \frac{\beta}{\alpha}. & (10)
\end{equation}

To understand this expression, each instant the individual remains infected, he or she infects a number of others equal to $\beta$ times the size of the susceptible population, which is approximately 1 since the infected individual is introduced into a fully susceptible population. The individual remains infected for an expected duration of $1/\alpha$.\(^8\) The subsequent analysis takes $R_0$ as the key exogenous parameter, capturing the disease’s infectiousness.\(^9\)

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, $S_\infty(Q) = \lim_{t \to \infty} S_t(Q)$.

The following series of lemmas, which characterize $S_t(Q)$ and $I_t(Q)$ for finite and limiting values of $t$, helps streamline the subsequent analysis. Many of the proofs are sketched in Martcheva (2015); online appendix A.1 provides full details.

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

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

**Lemma 3.** If $R_0 S_0(Q) \leq 1$, 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 $S_T(Q) = 1/R_0$, 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.** $R_0 S_\infty(Q) < 1$.

Intuitively, 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, the path of infections over the epidemic has two possible shapes: monotonically decreasing or hump shaped, expanding up to a peak and declining thereafter. The shape of the path hinges on the product $R_0 S_0(Q)$. Multiplying $R_0$, the expected number of secondary cases an infectious individual transmits in a fully susceptible population, by $S_0(Q)$, the proportion of susceptibles in the relevant

---

\(^6\)This model of vaccination is called *vaccination at recruitment* (Martcheva, 2015, sec. 9.2.1), distinct from *continuous vaccination* (Martcheva, 2015, sec. 9.2.2).

\(^7\)Logistical constraints would prevent such rapid vaccine rollout in practice, but the model may reasonably approximate an intensive vaccine campaign against COVID-19 or other epidemic disease.

\(^8\)To see this, note that the sole risk of exiting the infected state is recovery, with hazard $\lambda_S(t) = \alpha$. In Poisson duration models, the duration of a spell equals the reciprocal of the hazard, here $1/\lambda_S(t) = 1/\alpha$.

\(^9\)Estimates of $R_0$ vary considerably across diseases, from 1.1 for SARS (Chowell et al., 2003) at the low end to 16 to 18 for measles and pertussis at the high end (Anderson & May, 1991). Estimates of $R_0$ also vary across time and region.
population, yields the expected number of secondary cases an infectious individual transmits in the relevant population, called the effective reproductive number. If \(R_0S_0(Q) \leq 1\), there are fewer secondary infections than primary infections in the initial population, leading the infection rate to fall initially. Otherwise, the infection rate rises initially.

Lemma 5 says that the effective reproductive number cannot exceed 1 at the end of the epidemic. If \(R_0S_0(Q) > 1\), infections would be increasing, implying a growing, not waning, epidemic.

The term \(S_\infty(Q)\)—the proportion of people who remain healthy throughout the epidemic despite not being successfully immunized—plays a key role in the subsequent analysis, factoring into private and social benefits, thus determining equilibrium outcomes and the efficiency of these outcomes. While no closed-form solution exists for \(S_\infty(Q)\), the next lemma expresses it as an implicit function of other model parameters. The lemma also provides an expression for \(S_\infty(Q)\) in terms of the principal branch of the Lambert W function, here denoted \(L\).

**Lemma 6.** \(S_\infty(Q)\) satisfies
\[
\ln S_\infty(Q) - R_0S_\infty(Q) = \ln(\hat{S}_0 - \theta Q) - R_0(\hat{\theta}_0 + \hat{S}_0 - \theta Q)
\]
\[
(11)
\]
and can be written as
\[
S_\infty(Q) = \frac{1}{R_0} \left| L\left( - R_0(\hat{S}_0 - \theta Q)e^{-R_0(\hat{\theta}_0 + \hat{S}_0 - \theta Q)} \right) \right|.
\]
\[
(12)
\]
Comparative-static results can be obtained by applying the implicit function theorem to equation (11)—for instance,
\[
\frac{\partial S_\infty(Q)}{\partial Q} = \frac{\theta S_\infty(Q)}{S_0(Q)} \left[ \frac{R_0S_0(Q) - 1}{1 - R_0S_\infty(Q)} \right].
\]
\[
(13)
\]
Since its denominator is positive by lemma 5, the sign of equation (13) depends on whether the effective reproductive number, \(R_0S_0(Q)\), exceeds 1 initially. The immediate effect of an increase in \(Q\) is to move an individual from the currently susceptible to the vaccinated compartment. If \(R_0S_0(Q) \leq 1\), implying that the infection rate declines monotonically throughout the epidemic, this immediate effect carries through to a reduction in the final susceptible proportion, \(S_\infty(Q)\). If \(R_0S_0(Q) > 1\), implying that the infection rate initially increases, the reduction in current susceptibles has such a strong feedback effect, reducing epidemic growth, that the final susceptible proportion \(S_\infty(Q)\) increases despite the immediate reduction in susceptibles.

**B. Consumer Demand**

Consumers are homogeneous and risk neutral. Consistent with the short-run perspective adopted in this paper, assume that 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 \(MPB(Q) = \theta H \Phi(Q)\), where \(\Phi(Q)\) denotes the probability a susceptible contracts the disease during the epidemic.

To compute \(\Phi(Q)\), 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
\[
\Phi(Q) = \frac{1 - S_\infty(Q)}{S_0(Q)} = 1 - \frac{S_\infty(Q)}{S_0 - \theta Q},
\]
\[
(14)
\]
which lemma 4 guarantees is positive. Thus,
\[
MPB(Q) = \theta H \left[ 1 - \frac{S_\infty(Q)}{S_0 - \theta Q} \right].
\]
\[
(15)
\]
Differentiating, substituting from equation (13), and rearranging yields
\[
\frac{\partial MPB(Q)}{\partial Q} = \frac{-\theta R_0 S_\infty(Q) MPB(Q)}{S_0(Q)[1 - R_0 S_\infty(Q)]},
\]
\[
(16)
\]
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 equation (15) 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 equation (11) and solving for \(Q\) gives the following expression for demand: when some but not all consumers purchase:
\[
d(P) = \frac{1}{\theta H} \left[ \frac{\hat{S}_0 + \theta H}{P} \ln \left( 1 - \frac{P}{\theta H} + \hat{\theta}_0 \right) \right].
\]
\[
(17)
\]
Combining these facts yields the demand curve,

\[ D(P) = \begin{cases} 
0 & P > MPB(0) \\
d(P) & P \in [MPB(\bar{S}_0), MPB(0)] \\
\bar{S}_0 & P < MPB(\bar{S}_0) 
\end{cases} \] (18)

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

C. Firm Supply

We analyze two different market structures in the text: perfect competition and monopoly. Online appendix A.3 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 \( \Pi \) from date-0 sales.

By equation (14) and lemma 2, \( \Phi(Q) < 1 \), implying \( MPB(Q) < \theta H \) by equation (15). 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

\[ \frac{c}{\theta H} = \hat{c} < 1, \] (19)

introducing \( \hat{c} \) as shorthand notation to streamline subsequent expressions.

D. 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 those who never contract the disease:

\[ SB(Q) = H[1 - R_\infty(Q)] = H[S_\infty(Q) + \theta Q], \] (20)

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:

\[ W(Q) = SB(Q) - cQ. \] (21)

Marginal social benefit is the derivative \( MSB(Q) = \frac{\partial SB(Q)}{\partial Q} \). Differentiating equation (20), substituting from equations (13) to (15), and rearranging yields

\[ MSB(Q) = \frac{MPB(Q)}{1 - \mathcal{R}_0 S_\infty(Q)}. \] (22)

Let \( MEX(Q) = MSB(Q) - MPB(Q) \) denote the marginal externality from a vaccine course. Substituting from equation (22) yields

\[ MEX(Q) = \frac{\mathcal{R}_0 S_\infty(Q) MPB(Q)}{1 - \mathcal{R}_0 S_\infty(Q)}. \] (23)

Let \( Q^{**} \) denote the first-best quantity, maximizing \( W(Q) \). If \( Q^{**} \) is not a corner solution, involving either no vaccination or universal vaccination, it is an interior solution solving the social planner’s first-order condition \( MSB(Q^{**}) = c \).

III. Equilibrium

A. Perfect Competition

Equilibrium values of variables are distinguished with stars and a subscript indicating the relevant market structure. Under perfect competition, the equilibrium price is \( P^* = c \) and profit is \( \Pi^* = 0 \). 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 \( \mathcal{R}_0 \).

The table distinguishes three relevant cases corresponding to three intervals for \( \mathcal{R}_0 \). In case i, \( \mathcal{R}_0 \) is so low that no consumer finds it worthwhile to purchase the vaccine. The moderate values of \( \mathcal{R}_0 \) in case ii lead some but not all susceptibles to purchase. To compute the boundary value of \( \mathcal{R}_0 \) between cases i and ii denoted \( \mathcal{R}_0' \), note that equilibrium price \( P^*_c = c \) just choke off demand at this boundary. Setting \( d(c) = 0 \) in equation (17) and solving for \( \mathcal{R}_0 \) yields

\[ \mathcal{R}_0' = \frac{|\ln(1 - \hat{c})|}{1 - \hat{c} S_0}. \] (24)

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: \( Q^*_c = \hat{S}_0 = Q^{**} \). To compute the boundary value of \( \mathcal{R}_0 \) between cases ii and iii, denoted \( \mathcal{R}_0'' \), note that equilibrium price \( P^*_c = c \) just induces all susceptibles to purchase at this boundary. Setting \( d(c) = \hat{S}_0 \) in equation (17) and solving for \( \mathcal{R}_0 \) yields

\[ \mathcal{R}_0'' = \frac{|\ln(1 - \hat{c})|}{1 - (1 - \hat{c}) S_0}. \] (25)

To visualize how the variables in table 1 vary with \( \mathcal{R}_0 \), figure 1 graphs a selection of them as functions of \( \mathcal{R}_0 \). Focus for now on the dotted curves representing equilibrium under perfect competition. Vaccine quantity \( Q^*_c \), graphed in the first panel, rises throughout case ii from its value of 0 in case i to the first-best value \( Q^{**} \) in cases iii and iv. It is unsurprising that the equilibrium quantity is weakly increasing in the infectiousness 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 \( c \).
Consider the comparative-static effect of the weak change is strict. In cases iii and iv, the marginal ex-
case ii. Given that some consumers purchase in this case, consumers must be indifferent between purchasing and not, implying that the equilibrium price $P_c^* = c$ must exceed the entire marginal private benefit, implying $MPB_c^* = c$ over the entire interval.

Other variables display interesting nonmonotonicities. Cumulated infections over the epidemic $R_\infty(Q_c^*)$ initially increase in case i due to the epidemiological effects of the higher $R_0$. In case ii, when consumers begin purchasing vaccine, $R_\infty(Q_c^*)$ reverses course, sloping downward in $R_0$. The counterintuitive downward slope can be explained by the risk-compensation effect à la Peltzman (1975): the direct effect of an increase in infectiousness is more than offset by consumers’ behavioral response in the form of increased vaccine purchases. In cases iii and iv, $R_\infty(Q_c^*)$ again rises with $R_0$ because the direct effect of an increase in infectiousness cannot be offset by an increase in vaccine purchases given that all susceptibles are vaccinated. The marginal externality $MEX_c^*$ exhibits an even more complex nonmonotonic pattern. The interplay between increasing infectiousness and increasing vaccine quantity generates two local maxima in the figure, with the global maximum occurring at the boundary between cases i and ii.

Proposition 1 summarizes the comparative-static effects of an increase in $R_0$ on the steady-state equilibrium under perfect competition, showing that the observations from figure 1 are quite general. Online appendix A.1 provides proofs for results not obvious from table 1.

**Proposition 1.** Consider the comparative-static effect of $R_0$ on equilibrium variables under perfect competition:

- Price and industry profit are constant: $P_c^* = c$ and $\Pi_c^* = 0$, respectively.
- Quantity $Q_c^*$ and marginal private benefit $MPB_c^*$ are weakly increasing in $R_0$.
- Welfare $W_c^*$ is weakly decreasing in $R_0$.
- Cumulated infection, $R_\infty(Q_c^*)$, attains a single interior local maximum in $R_0$, which is a global maximum if and only if $\tilde{c} > 1 - 0$.
- For the marginal social benefit $MSB_c^*$ and marginal externality $MEX_c^*$, each attains no more than two interior local maxima in $R_0$, one of which is a global maximum.

For each $Q_c^*$, $MPB_c^*$, and $W_c^*$, there exists a nonempty interval of $R_0$ such that the weak change is strict.

Note that the Peltzman effect identified for cumulated infections, whereby an increase in $R_0$ leads to an equilibrium reduction in $R_\infty(Q_c^*)$, does not extend over the whole range of $R_0$ but just over case ii. Note further that the reductions in infections in case ii do not translate into an increase in welfare, which is weakly decreasing for all $R_0$. Increased consumer spending on vaccines more than offsets the reduction in cumulated infections.

B. Monopoly

Since price equals cost under perfect competition, but a monopoly charges a markup above cost, price is weakly higher under monopoly and quantity weakly lower. It follows that in case i, in which $Q_c^* = 0$, we have $Q_m^* = 0$. 

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

Other variables display interesting nonmonotonicities. Cumulated infections over the epidemic $R_\infty(Q_c^*)$ initially increase in case i due to the epidemiological effects of the higher $R_0$. In case ii, when consumers begin purchasing vaccine, $R_\infty(Q_c^*)$ reverses course, sloping downward in $R_0$. The counterintuitive downward slope can be explained by the risk-compensation effect à la Peltzman (1975): the direct effect of an increase in infectiousness is more than offset by consumers’ behavioral response in the form of increased vaccine purchases. In cases iii and iv, $R_\infty(Q_c^*)$ again rises with $R_0$ because the direct effect of an increase in infectiousness cannot be offset by an increase in vaccine purchases given that all susceptibles are vaccinated. The marginal externality $MEX_c^*$ exhibits an even more complex nonmonotonic pattern. The interplay between increasing infectiousness and increasing vaccine quantity generates two local maxima in the figure, with the global maximum occurring at the boundary between cases i and ii.

Proposition 1 summarizes the comparative-static effects of an increase in $R_0$ on the steady-state equilibrium under perfect competition, showing that the observations from figure 1 are quite general. Online appendix A.1 provides proofs for results not obvious from table 1.

**Proposition 1.** Consider the comparative-static effect of $R_0$ on equilibrium variables under perfect competition:

- Price and industry profit are constant: $P_c^* = c$ and $\Pi_c^* = 0$, respectively.
- Quantity $Q_c^*$ and marginal private benefit $MPB_c^*$ are weakly increasing in $R_0$.
- Welfare $W_c^*$ is weakly decreasing in $R_0$.
- Cumulated infection, $R_\infty(Q_c^*)$, attains a single interior local maximum in $R_0$, which is a global maximum if and only if $\tilde{c} > 1 - 0$.
- For the marginal social benefit $MSB_c^*$ and marginal externality $MEX_c^*$, each attains no more than two interior local maxima in $R_0$, one of which is a global maximum.

For each $Q_c^*$, $MPB_c^*$, and $W_c^*$, there exists a nonempty interval of $R_0$ such that the weak change is strict.

Note that the Peltzman effect identified for cumulated infections, whereby an increase in $R_0$ leads to an equilibrium reduction in $R_\infty(Q_c^*)$, does not extend over the whole range of $R_0$ but just over case ii. Note further that the reductions in infections in case ii do not translate into an increase in welfare, which is weakly decreasing for all $R_0$. Increased consumer spending on vaccines more than offsets the reduction in cumulated infections.

B. Monopoly

Since price equals cost under perfect competition, but a monopoly charges a markup above cost, price is weakly higher under monopoly and quantity weakly lower. It follows that in case i, in which $Q_c^* = 0$, we have $Q_m^* = 0$. 

---

Mathematically, to maintain the constant marginal private benefit ($MPB_c^* = c$) observed throughout case ii, the increase in infectiousness $R_0$ must be offset by a reduction in infections to maintain a constant probability of contracting the disease.
Case i is thus trivially identical across perfect competition and monopoly. In the remaining cases, perfectly competitive firms are able to make positive sales at price $c$. By continuity, the monopoly can make positive sales at some small markup above $c$, implying $Q_m^* > 0$ for $R_0$ in cases ii and above.

To solve for $Q_m^*$ in these other cases, the monopoly’s maximization problem can be transformed so that the choice variable is quantity rather than price. The monopoly optimally sets a price to extract the entire private benefit of the marginal consumer, leading to inverse demand $P(Q) = MPB(Q)$. The monopoly chooses $Q$ to maximize $[MPB(Q) - c]Q$ subject to $Q \leq \tilde{S}_0$, a constrained maximization problem that can be solved using the Kuhn-Tucker method. Denote the monopoly’s marginal revenue by $MR(Q) = \partial[MPB(Q)Q]/\partial Q$. Using equation (16), one can show that the preceding derivative equals

$$MR(Q) = MPB(Q) \left\{ 1 - \frac{\theta R_0 Q [1 - \Phi(Q)]}{1 - R_0 S_\infty(Q)} \right\}. \tag{26}$$

According to standard Kuhn-Tucker conditions, the solution is an interior optimum satisfying the textbook condition $MR(Q_m^*) = c$ unless $MR(\tilde{S}_0) \geq c$, in which case the solution is the corner, $Q_m^* = \tilde{S}_0$. The following proposition records this solution for equilibrium monopoly output and compares it to output under perfect competition.

**Proposition 2.** For sufficiently low $R_0$, namely, for all $R_0 \leq R_0^*$, there is no output in equilibrium under either monopoly or perfect competition: $Q_m^* = Q^*_c = 0$. For sufficiently high $R_0$, namely, for all $R_0$ satisfying $MR(\tilde{S}_0) \geq c$, equilibrium output attains the first best—which for these values of $R_0$ involves universal vaccination of susceptibles—under both monopoly and perfect competition: $Q_m^* = Q^*_c = Q^{**} = \tilde{S}_0$. Otherwise, equilibrium monopoly output is an interior value $Q_m^* \in (0, \tilde{S}_0)$ satisfying $MR(Q_m^*) = c$ and is strictly lower than output under perfect competition: $Q_m^* < Q^*_c$.

The proof in online appendix A.1 fills in details omitted from the sketch preceding the proposition, including verifying that the condition $MR(\tilde{S}_0) \geq c$ is satisfied for sufficiently high $R_0$.

The values of $Q_m^*$ and other equilibrium variables are cataloged in table 2. The entry for $Q_m^*$ (and by extension, the other equilibrium values expressed in terms of $Q_m^*$) are not provided in analytic form, let alone in closed form, in cases ii and iii. This need not preclude definitive comparative-statics results; one could apply the implicit function theorem to the condition $MR(Q_m^*) = c$ to determine how $Q_m^*$ changes with $R_0$ in those cases. However, this approach still does not deliver a definitive sign. We can be sure that $Q_m^*$ increases in $R_0$ for some $R_0$ in cases ii and iii—since $Q_m^*$ must rise from 0 to the first-best quantity $\tilde{S}_0$ somewhere in that set by continuity—but we have not been able to rule out the possibility that the monopoly responds to an increase in $R_0$ 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 some equilibrium variables, reported in the next proposition, proved in online appendix A.1.

**Proposition 3.** Consider the comparative-static effect of $R_0$ on equilibrium variables under monopoly:

- Monopoly profit $\Pi_m^*$ is weakly increasing in $R_0$. For $R_0 \leq R_0^*$, $\Pi_m^* = 0$. For $R_0 > R_0^*$, $\Pi_m^*$ is positive and strictly increasing in $R_0$.

- Cumulative infections $R_\infty(Q_m^*)$ attains one or more interior local maxima in $R_0$, one of which is a global maximum if $c \geq 1 - \theta$.

- Both marginal social benefit $MSB_m^*$ and marginal externality $MEX_m^*$ attain an interior global maximum in $R_0$.

Equilibrium values of selected variables under monopoly are graphed as functions of $R_0$ as the solid curves in figure 1.
The two market structures overlap in case i, neither generating any vaccine output. The two market structures overlap again in case iv, both generating the first-best quantity $Q^{**} = \tilde{S}_0$. In between—in cases ii and iii—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 large gap between the dotted and solid curves for intermediate values of $R_0$ suggests that the distortion arising from the monopoly’s exercise of its market power is worst for moderate levels of infectiousness. The marginal externality can be considerably higher under monopoly for some $R_0$ but can be slightly lower for some $R_0$ as the lower monopoly output generates higher marginal private benefit, leaving less residual externality.

The graph of $W^*$ under monopoly illustrates the remarkable possibility that increasing $R_0$ can increase welfare, impossible under perfect competition according to proposition 1. Under monopoly, not only do consumers fail to consider the external benefit their vaccination provides other consumers, but the monopoly compounds this by placing negative value on consumption to the extent it reduces others’ willingness to pay for a vaccine. An increase in $R_0$ can mitigate this compound underconsumption problem, providing such a large indirect benefit that it swamps the direct harm from increased infectiousness, leading to an increase in social welfare.

### IV. Government Subsidies

We have seen that free riding can lead to inefficiently low vaccination 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 course subsidy $G \geq 0$ at the outset of the game. According to standard public finance results, the economic incidence is the same whether consumers or firms are the statutory target of a tax or subsidy (Fullerton & Metcalf, 2002). We adopt the accounting convention that $G$ is paid to firms, in which case the subsidy is equivalent to a reduction in firms’ marginal cost from $c$ to $c - G$. Since social welfare is maximized by the first-best quantity $Q^{**}$, the first-best subsidy $G^{**}$ is that implementing $Q^{**}$.

The set of $R_0$ satisfying $MR(S_0) \geq c$ need not form an interval but does include all sufficiently high $R_0$. Any value $F^m_R \geq c$ is consistent with zero sales in equilibrium.

### Table 2.—Equilibrium Variables under Monopoly as Functions of $R_0$

| Case | Variable | (i) $R_0 \in (0, R^*_0]$ | (ii, iii) $R_0 > R^*_0$ but $MR(S_0) < c$ | (iv) $R_0$ satisfies $MR(S_0) \geq c$ |
|------|----------|--------------------------|--------------------------------|--------------------------------|
| $P^*_m$ | $\theta H \Phi(Q^*_m)$ | $\theta H \Phi(S_0)$ | $\theta H \Phi(S_0)$ |
| $Q^*_m$ | 0 | Solution to $MR(Q^*_m) = c$ | $S_0$ |
| $\Pi^*_m$ | $\theta H \Phi(0)$ | $S_0$ |
| $MR^*_m$ | $\theta H \Phi(0)$ | $S_0$ |
| $MSR^*_m$ | $\frac{\theta H \Phi(0) - MR^*_m}{R_0S_0(0)}$ | $\frac{\theta H \Phi(0) - MR^*_m}{R_0S_0(0)}$ | $\frac{\theta H \Phi(0) - MR^*_m}{R_0S_0(0)}$ |
| $W^*_m$ | $HS_0(0)$ | $H[S_0(0) + \theta(1 - \epsilon)Q^*_m]$ | $H[S_0(0) + \theta(1 - \epsilon)S_0]$ |

Equation (12) provides a formula for computing $S_0(0), S_m(0)$, and $S_m(Q^*_m)$. Equation (14) provides formulas for computing $\Phi(0), \Phi(S_0)$, and $\Phi(S^*_0)$. The distinction between cases ii and iii in the middle column, relevant for perfect competition in previous table, is irrelevant for monopoly here. As indicated in equation (26), $MR(S_0)$ is a function of $R_0$, though for brevity, $R_0$ is not included in the argument list for $MR$. It immediately extends to perfect competition or any market structure. Since the marginal vaccine externality has lexicographic preferences over welfare and expenditure savings.

It is straightforward to establish a set of broad results for any market structure. Since the marginal vaccine externality is nonnegative by equation (23), equilibrium output $Q^*$ never exceeds the first best $Q^{**}$. If $Q^{**} = 0$, then $Q^* = Q^{**} = 0$ as well, implying $G^{**} = 0$ since the first best can be achieved without a subsidy. The proof of the next proposition shows that $Q^{**} = 0$ for all $R_0$ in a neighborhood above 0, implying that $G^{**} = 0$ in this neighborhood for any market structure.

We can also draw broad conclusions about the optimal subsidy for high values of $R_0$. By proposition 2, $Q^*_m = \tilde{S}_0$ for sufficiently high values of $R_0$. For such $R_0$, $\tilde{S}_0 = Q^*_m = Q^{**} \leq \tilde{S}_0$, implying $Q^*_m = Q^{**}$, in turn implying $G^{**} = 0$ since the first best can be achieved without a subsidy under monopoly. The result that $G^{**} = 0$ for sufficiently high $R_0$ immediately extends to perfect competition or any market structure involving weakly higher output than monopoly.

Having established that $G^{**} = 0$ for intervals of low and high values of $R_0$ for general market structures, if it can be shown that $G^{**} > 0$ for some intermediate value of $R_0$, it is
immediate that \( G^{**} \) is nonmonotonic, attaining a global maximum for some interior \( R_0 \in (0, \infty) \) as the next proposition states. The proof provided in online appendix A.1 fills in this and other omitted details.

**Proposition 4.** For monopoly—or any market structure involving weakly higher output including perfect competition—the optimal government subsidy equals 0 for sufficiently low and sufficiently high \( R_0 \) and attains an interior global maximum in \( R_0 \).

As proposition 4 indicates, the optimal subsidy is not monotonically increasing in \( R_0 \) as might be inferred based solely on epidemiological considerations but is maximized for an interior value of \( R_0 \). The difficulty in addressing a disease depends not only on its infectiousness but also on consumers’ response to this infectiousness. Free riding on the vaccination of others disappears with extremely infectious diseases; moderately infectious diseases provide consumers more leeway to free-ride, requiring a higher optimal subsidy to address.

We conclude the section with a more precise precise characterization of the optimal subsidy under perfect competition and monopoly provided by the next proposition, proved in online appendix A.1.

**Proposition 5.** The optimal government subsidies under perfect competition and monopoly depend on the first-best output, \( Q^{**} \):

- If the first best involves no output, that is, \( Q^{**} = 0 \), then no subsidy is needed under either market structure: \( G^{**}_c = G^{**}_m = 0 \).

- If the first best involves an interior output level, \( Q^{**} \in (0, \hat{S}_0) \), then the optimal subsidy under perfect competition corrects for the externality at the target output: \( G^{**}_c = \text{MEX}(Q^{**}) \). The optimal subsidy under monopoly is related but is scaled up to offset the monopoly’s only partial pass-through: \( G^{**}_m = \text{MEX}(Q^{**}) \hat{S}_0/(\hat{S}_0 - \theta Q^{**}) \).

- If the first best involves universal vaccination of susceptibles, \( Q^{**} = \hat{S}_0 \), then the optimal subsidy under perfect competition bridges the gap, if any, between consumers’ marginal private benefit under universal vaccination and marginal cost, \( c \): \( G^{**}_c = \max[0, c - \text{MPB}(\hat{S}_0)] \). The optimal subsidy under monopoly is related but needs to bridge a larger gap due to the monopoly markup:

\[
G^{**}_m = \max \left[ 0, c - \text{MPB}(\hat{S}_0) + \frac{\theta}{1 - \theta} \text{MEX}(\hat{S}_0) \right].
\]

(27)

Across all cases, the optimal subsidy is weakly higher under monopoly than perfect competition, \( G^{**}_m \geq G^{**}_c \), strictly so if \( Q^{**}_m \in (0, \hat{S}_0) \).

**V. Universal Vaccination**

Under both market structures, equilibrium attains universal vaccination of susceptibles for a nonempty set of parameters. Two reservoirs of infection motivate the marginal consumer to purchase at a positive price even when all other consumers also purchase. With an imperfectly effective vaccine (\( \theta < 1 \)), some vaccinated consumers remain susceptible and able to transmit the disease to others. Even in the limit of perfect efficacy \( \theta \uparrow 1 \), however, the \( I_0 \) initially infected individuals remain a reservoir. Previous game-theoretic analyses finding that a perfectly effective vaccine would never be universally purchased at a positive price (Geoffard & Philipson, 1997; May, 2000; Bauch & Earn, 2004) omitted this feature of the SIR model.

It is obvious that equilibrium under perfect competition must attain universal vaccination if the disease is infectious enough. Even a small reservoir of infecteds \( I_0 \), when combined with a sufficiently high \( R_0 \), generates high enough infection risk to motivate the marginal consumer to purchase at any fixed \( c < H \). It is less obvious that universal vaccination is attained in a monopoly equilibrium. Given that its monopoly price is endogenous, not fixed, the monopoly might respond to an increase in \( R_0 \) by raising price, keeping output short of universal vaccination. For sufficiently high \( R_0 \), the marginal consumer is almost certain to contract the disease from the reservoir of \( I_0 \) infecteds irrespective of how many susceptibles are vaccinated. The monopoly serves all consumers at a price approaching the marginal private benefit of being protected against certain infection, leaving little room for any further price increase without losing most customers.

The next proposition, proved in online appendix A.1, provides a simple necessary and sufficient condition for universal vaccination with a perfectly effective vaccine to obtain in equilibrium under each market structure.

**Proposition 6.** In the limit of a perfectly effective vaccine (\( \theta \uparrow 1 \)), universal vaccination of susceptibles is attained in equilibrium under perfect competition if and only if \( 1 - e^{-R_0 \hat{I}_0} > \check{c} \) and under monopoly if and only if \( 1 - e^{-R_0 \hat{I}_0}(1 - R_0 \hat{S}_0 e^{-R_0 \hat{I}_0}) > \check{c} \).

It can be shown that both conditions hold for sufficiently high \( R_0 \): the factors on the left-hand side of both conditions are equal 1 in the limit \( R_0 \uparrow \infty \) and 1 > \( \check{c} \) by assumption (19). It is also obvious that neither condition holds for any finite \( R_0 \) when \( I_0 = 0 \), reflecting the fact that an initial stock of infecteds is required to generate demand for a perfectly effective vaccine under a universal vaccination program.

**VI. 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 to 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 \( MSB(\theta) \) is increasing in \( \theta \). Differentiating equation (22), substituting from equation (13), and rearranging yields

\[
\frac{\partial MSB(\theta)}{\partial \theta} = \frac{6^{\theta}H R_0 S_\infty(\theta) S_0(\theta) \Phi(\theta)}{[1 - R_0 S_\infty(\theta)]^3} \times (R_0[S_0(\theta) + S_\infty(\theta)] - 2). \tag{28}
\]

All the factors on the right-hand side are definitively positive by lemma 5 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 7.** The \( Q \)th unit of vaccine exhibits increasing social returns if and only if

\[
R_0 \left[ \frac{S_0(\theta) + S_\infty(\theta)}{2} \right] > 1. \tag{29}
\]

Earlier, we identified the inequality \( R_0 S_0(\theta) > 1 \) as necessary and sufficient for the epidemic to grow rather than decline from the start. Condition (29) is more stringent. Instead of requiring the initial value of the effective reproductive number, \( R_0 S_0(\theta) \), to exceed 1, it requires the average of the initial value \( R_0 S_0(\theta) \) and the final value \( R_0 S_\infty(\theta) \) to exceed 1. By lemma 4, \( R_0 S_\infty(\theta) < R_0 S_0(\theta) \). Proposition 7 can thus be interpreted as saying that unit \( \theta \) of the vaccine exhibits increasing social returns if the potential not just for immediate but for sustained epidemic expansion is sufficiently high.

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 online appendix A.1 as a straightforward corollary of proposition 7.

**Proposition 8.** The vaccine exhibits increasing social returns initially—at an output level of \( Q = 0 \) if \( R_0 S_0(\theta) \geq 2 \). The vaccine exhibits increasing social returns everywhere—at all output levels \( Q \in (0, S_0) \) if \( R_0 S_0(\theta) \geq 2/(1 - \theta) \).

According to proposition 8, if a federal authority has access to only 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 \( R_0 S_0(\theta) > 2 \). If, for example, \( S_0 = 0.8 \) in each state, then concentrating the vaccine would be efficient for any \( R_0 > 2.5 \). If the more stringent condition \( R_0 S_0(\theta) > 2/(1 - \theta) \) 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 \( R_0 \) would need to be cleared for concentrating vaccines in one state to be more efficient than serving high-value consumers everywhere first.

**VII. 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) lists 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’ privacy relative to drug molecules, to the scale often needed for vaccine intellectual property to disease-risk information when making drug purchases (the focus of that paper).

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 prefer to develop a drug that does not have this demand-reducing public-good feature than a similarly effective vaccine.

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 delicate since at equal marginal production costs \( e \), 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: \( c = 0 \). Assume the drug is effective with probability \( \theta \). 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 previously computed monopoly profit and welfare from a vaccine, respectively, \( \Pi_{md}^* \) and \( W_{md}^* \), it remains to compute the analogous variables for a drug, respectively, \( \Pi_{md}^* \) 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 afterward, yielding drug profit

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

(30)

To compute equilibrium welfare with a drug, the \( \hat{I}_0 \) individuals infected initially along with the \( S_0 - S_{\infty}(0) \) infected later obtain health benefit \( H \) with probability \( \theta \) 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:

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

(31)

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 online appendix A.1. The proposition uses the notations \( \Delta \Pi_{md}^* = \Pi_{md}^* - \Pi_{mv}^* \) and \( \Delta W_{md}^* = W_{md}^* - W_{mv}^* \) for differences between equilibrium variables for the two products and \( \Delta W_{vs}^* = W_{vs}^* - W_{vs}^* \) for difference between first-best welfare.

**Proposition 9.** Suppose \( c = 0 \). For all \( \mathcal{R}_0 > 0 \), monopoly profit is strictly higher from a drug than vaccine: \( \Delta \Pi_{md}^* > 0 \). The profit advantage from a drug \( \Delta \Pi_{md}^* \) approaches its lowest value, \( \inf_{\mathcal{R}_0 > 0} \Delta \Pi_{md}^* = \theta H \hat{I}_0 \), in the limits of extremely low and extremely high \( \mathcal{R}_0 \) and attains an interior global maximum in \( \mathcal{R}_0 \). Welfare is higher with a drug than a vaccine for extremely low and extremely high \( \mathcal{R}_0 \): \( \lim_{\mathcal{R}_0 \downarrow 0} \Delta W_{md}^* > 0 \) and \( \lim_{\mathcal{R}_0 \uparrow \infty} \Delta W_{md}^* > 0 \). However, there exist parameters for which welfare is higher with a vaccine: \( \Delta W_{md}^* < 0 \).

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 with an infinitely infectious disease because susceptibles are certain to contract the disease, if no one else, from the \( \hat{I}_0 \) initially infected. 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.

**VIII. COVID-19 Calibrations**

This section provides a calibration using parameters drawn from the COVID-19 pandemic. The calibration is meant as an illustration, not a forecast. Our model is too stylized along many dimensions 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-economic considerations lead real-world vaccine markets to depart from our theoretical construct of firms selling directly to individual consumers without third-party funding.

Based on a meta-analysis of studies of the ancestral strain of COVID-19 (Liu et al., 2020), we set \( \mathcal{R}_0 = 2.8 \). We take estimates of needed population 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-19 vaccines. We use estimates from U.K. government agencies, which provide some of the best estimates for a developed country then available. Based on the U.K. Office for National Statistics (2020), we take the infected proportion at that time to be \( \hat{I}_0 = 0.19\% \) and the recovered proportion 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-19 infection (including both symptomatic and asymptomatic). For rescaled cost, \( \hat{c} = c/\theta H \), we take the limiting case of a costless vaccine, \( \hat{c} \downarrow 0 \), reflecting the low cost \( c \) for existing vaccines, especially in comparison to the potential disease harm \( H \), as documented further in online appendix A.4.

These parameters put us in case iii of tables 1 and 2, in which perfect competition attains the first best of universal vaccination but monopoly does not. Using numerical methods to compute \( S_{\infty}(Q) \) in equation (12) and to optimize monopoly profit, we find that the monopoly price is set to 49% of the harm from contracting the disease. At this price, only 51% of susceptible consumers buy, generating welfare equal to 59% of the available health benefit. The optimal
subsidy required to generate the first best under monopoly is enormous—over three times the equilibrium monopoly price.

The effective reproductive number, $R_0S_0 = 2.6$, surpasses the threshold of 2 sufficient for initially increasing social returns according to proposition 8 but not the threshold for everywhere increasing social returns, which equals $2/(1 - \theta) = 10$ given calibrated efficacy. Examining equation (29) for a range of quantities, one can determine that increasing returns persist through output equal to 63% 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 81% of the population of one state. Larger stockpiles than this are more efficiently divided equally between the states. Overall, the results suggest that social returns to COVID vaccines can be strongly increasing.

For the calibrated parameters, monopoly profit and welfare are both higher with a drug than a vaccine (that is, $\Delta \Pi_m > 0$ and $\Delta W_m > 0$). The monopoly's bias toward a drug thus does not lead to a distortion in the calibration.

To measure the sensitivity of the outcomes to infectiousness, we repeat the calibration holding all parameters constant at their October 2020 levels except $R_0$, replaced with its higher value for the delta variant: $R_0 = 5.1$ according to Liu and Rocklöv's (2021) meta-analysis. Monopoly quantity increases to 74% of the susceptible population, despite a 28% increase in the monopoly price. The higher vaccination rate leads to a 6% increase in equilibrium monopoly welfare in the delta calibration compared to that for ancestral COVID despite delta's greater infectiousness. That welfare may rise with infectiousness due to the response of economic agents appears to be not just a theoretical curiosity but may hold for realistic parameters. The increase in infectiousness reduces equilibrium welfare under perfect competition since the vaccination rate is already as high as possible when $R_0 = 2.8$, so the increase to $R_0 = 5.1$ results in a direct increase in disease burden.

**IX. Conclusion**

We analyzed the market for technologies that, by protecting individuals against an infectious disease, reduce transmission to others, a positive externality. Though the analysis applies to a variety of technologies such as circumcision, bed nets, and social distancing, the discussion focused on vaccines for concreteness. Vaccines (and most of the other technologies we noted) 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 consumption and production decisions.

Such distortions and policy correctives are the focus of this paper. To study them, we constructed a theoretical model of the vaccine market involving economic agents basing their consumption and production decisions on rational expectations of the disease’s evolution consistent with a standard SIR epidemiological model. Within that general framework, we made specific modeling choices to suit an intensive vaccine campaign against an epidemic disease such as COVID-19 expected to wane before generations turn over. We pursued a comprehensive account of equilibrium variables such as price, quantity, profit, and welfare across a variety of market structures ranging from perfect competition to monopoly and studied how those variables change in response to an increase in the infectiousness of the disease as measured by $R_0$, among other parameters. Since our comparative-statics results are derived in a model where sales are made on a private market without government intervention, they are perforce counterfactual for real-world vaccine markets in which the potential for severe market failures and lost lives leads policymakers to intervene. Understanding how markets perform in the counterfactual absence of intervention, however, is a useful step toward characterizing optimal interventions, as we seek to do.

Perhaps the variable of most interest was the equilibrium marginal externality. Across the range of market structures studied, we found that the equilibrium marginal externality peaks for intermediate rather than extreme values of $R_0$. For low levels of $R_0$, one consumer’s vaccination provides little benefit to others because there is little chance the consumer would have infected them. For high levels of $R_0$, one consumer’s vaccination provides little benefit to others since they will most likely contract it from a different source anyway. Other outcome variables also peak for intermediate values of $R_0$—outcome variables including $G$ (the minimal subsidy necessary to obtain the first-best vaccine quantity) and $\Delta \Pi_m$ (extra monopoly profit from a drug that does not exert the epidemiological externality compared to a vaccine that does). Moderately infectious diseases may exhibit the greatest distortions and be prime targets for subsidy.

Across the range of market structures studied, universal vaccination of susceptible consumers is obtained in equilibrium for sufficiently high $R_0$. This is true even for a perfectly effective vaccine, contrasting some impossibility results in the previous literature. The key to this result is the presence of the $I_0$ infected individuals at vaccine rollout. Even if all other susceptibles are successfully vaccinated, the threat of contracting the disease from the $I_0$ infected induces the marginal consumer to purchase the vaccine at a positive price. In the limit of an arbitrarily infectious disease, free riding is eliminated since the risk of contracting the disease from even a small $I_0$ approaches 1. The presence of the $I_0$ infecteds also raises the possibility that a vaccine with a positive epidemiological externality can be welfare-dominated.

---

12 The second calibration is a thought experiment, not meant to reflect the situation when delta emerged, which involved several coexisting strains, a vaccine campaign already underway optimized against an earlier strain, and different proportions of susceptibles, infected, and recovered individuals.
by a drug without it: if the externality is small, welfare may be driven by the advantage of the drug in treating the $I_0$ infected for whom the vaccine arrives too late to help (assuming the vaccine must be administered prior to infection to be effective).

We derived simple sufficient conditions under which vaccination exhibits increasing social returns: social returns are initially increasing if $R_0 S_0 \geq 2$ and everywhere increasing if $R_0 S_0 \geq 2/(1 - \theta)$. If the first condition holds, a small supply is more efficiently 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 the supply. These stark implications for concentrating supplies hinge on the homogeneity of consumers in the model but raise the possibility that equitable allocation can lead to inefficiency.

In our calibration to the ancestral strain of COVID-19, we found that a competitively supplied vaccine would attain the first best of universal vaccination, but a monopoly would not. A monopoly—at least one unconstrained by public repugnance (à la Roth, 2007) against “profiting during a pandemic”—sets such a high price that only about half of susceptibles buy. Correcting this distortion requires an enormous subsidy, equal to over three times the equilibrium monopoly price.

Such a subsidy is likely to be prohibitive in many practical settings, pointing to the appeal of an alternative policy—bulk purchases negotiated by the government on behalf of consumers—that could achieve the first best at a much lower expenditure level. In fact, negotiated bulk purchases were used for COVID-19 vaccines as well as vaccines for many childhood diseases. Our results for equilibrium on the private market remain relevant if, following Kremer and Snyder (2015, section IV.C), one assumes the private market provides the threat point for Nash bargaining over the bulk purchase.

The COVID-19 calibration exhibited increasing social returns through a substantial range (63%) of the susceptible population. The presence of strongly 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 by the benefit of vaccinating vulnerable subpopulations in every country.

Comparing the calibration for the ancestral COVID strain to a second calibration for the delta variant, doubling infectiousness raises equilibrium welfare under monopoly by inducing more consumers to purchase even at the higher monopoly price, offsetting the direct increase in disease burden. Equilibrium welfare falls under perfect competition since the universal vaccination already attained with the less infectious ancestral strain leaves no room for a further increase in the vaccination rate to offset the increase in disease burden.

Key results—including that moderately rather than severely infectious diseases may be prime targets for subsidy—are robust to a variety of modeling alternatives. The results derived here for perfect competition and monopoly are extended in online appendix A.3 to Cournot competition among $n$ firms, nesting these other market structures as special cases. The assumption of homogeneous consumers is relaxed in online appendix A.4, which allows for heterogeneity in consumer harm, $H_x$, as an illustrative example. Our companion paper (Goodkin-Gold et al., 2022) maintains the SIR framework but adopts alternative modeling assumptions suited to an endemic disease against which new cohorts are continuously vaccinated in the steady state. The analysis in that paper is relevant to diseases such as measles, HIV, and even COVID-19 if continued emergence of new variants leads it to persist in the population over the long run. Despite mathematical differences—unlike here, steady states in the companion paper have simple closed-form expressions—the results are remarkably similar, down to the shape of the graphs of outcome variables against $R_0$, which resemble those in figure 1.

REFERENCES

Adida, Elodie, Debabrata Dey, and Hamed Mamani, “Operational Issues and Network Effects in Vaccine Markets,” European Journal of Operational Research 231:2 (2013), 414–427. 10.1016/j.ejor.2013.05.034

Acemoglu, Daron, Ali Makhdoomi, Azaraksh Malekian, and Asuman Ozdaglar, “Testing, Voluntary Social Distancing and the Spread of an Infection,” NBER working paper 27483 (2020).

Acemoglu, Daron, Victor Chernozhukov, Ivan Werning, and Michael D. Whinston, “Optimal Targeted Lockdowns in a Multigroup SIR Model,” American Economic Review: Insights 3:4 (2021), 487–502. 10.1257/aeri.20200590

Althouse, Benjamin M., Theodore C. Bergstrom, and Carl T. Bergstrom, “A Public Choice Framework for Controlling Transmissible and Evolving Diseases,” Proceedings of the National Academy of Sciences 107:suppl. 1 (2010), 1696. 10.1073/pnas.0906078107

Alvarez, Fernando, David Argente, and Francesco Lippi, “A Simple Planning Problem for COVID-19 Lock-Down, Testing, and Tracing,” American Economic Review: Insights 3:3 (2021), 367–382. 10.1257/aeri.20200201

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

Anderson, Soren T., Ramanan Laxminarayan, and Stephen W. Salant, “Diversify of Focus? Spending to Combat Infectious Diseases When Budgets Are Tight,” Journal of Health Economics 31:4 (2012), 658–675.

Atkeson, Andrew, Karen Kopecky, and Tao Zha, “Estimating and Forecasting Disease Scenarios for COVID-19 with an SIR Model,” NBER working paper 27335 (2020).

——— “Behavior and the Transmission of COVID-19,” American Economic Association Papers and Proceedings 111 (2021), 356–360.

Avery, Christopher, “A Simple Model of Social Distancing and Vaccination,” NBER working paper 29463 (2021).

Avery, Christopher, William Bossert, Adam Clark, Glenn Ellison, and Sara Ellison, “An Economist’s Guide to Epidemiology Models of Infectious Disease,” Journal of Economic Perspectives 34:4 (2020), 79–104.

Bauch, Chris T., and David J. D. Earn, “Vaccination and the Theory of Games,” Proceedings of the National Academy of Sciences 101:36 (2004), 13391–13394. 10.1073/pnas.0403823101

Boulou, Bryan, Tejwant Datta, and Robert Goldfarb, “Vaccination Externalities,” B.E. Journal of Economic Analysis and Policy 7:1 (2007), art. 23. 10.2202/1935-1682.1487

Brito, Dagobert L., Eytan Sheshinski, and Michael D. Intrilligator, “Externalities and Compulsory Vaccination,” Journal of Public Economics 45:1 (1991), 69–90. 10.1016/0047-2219(91)90048-7

Buckner, Jack H., Gerardo Chevall, and Michael R. Springborn, “Dynamic Prioritization of COVID-19 Vaccines When Social
