SOCIAL DISTANCING, VACCINATION AND THE PARADOXICAL OPTIMALITY
OF AN ENDEMIC EQUILIBRIUM

Andrew B. Abel
Stavros Panageas

Social Distancing, Vaccination and the Paradoxical Optimality of an Endemic Equilibrium
Andrew B. Abel and Stavros Panageas
NBER Working Paper No. 27742
August 2020, Revised April 2021
JEL No. C61,E61,I18

ABSTRACT

We analyze the impact of public health policy on the spread of a disease using a version of the SIR model that includes vital statistics, waning immunity, and vaccination. This model is rich enough to accommodate endemic steady states and disease-free steady states. We choose social distancing and vaccines to maximize an objective function that penalizes lost output resulting from social distancing, deaths resulting from the disease, and the cost of vaccination. Surprisingly, even though a disease-free equilibrium is attainable, optimal policy leads to an endemic steady state, though with a small number of deaths and negligible loss of output.

Andrew B. Abel
The Wharton School
University of Pennsylvania
2315 Steinberg Hall - Dietrich Hall
Philadelphia, PA 19104-6367
and NBER
abel@wharton.upenn.edu

Stavros Panageas
Anderson School of Management
University of California, Los Angeles
110 Westwood Plaza
Los Angeles, CA 90095-1481
and NBER
stavros.panageas@anderson.ucla.edu

"The public doesn't understand all about vaccines ... including that this disease may, even with vaccines, become endemic." Dr. David Heymann, WHO, Nov 12, 2020

"I doubt we are going to eradicate this. I think we need to plan that this is something we may need to maintain control over chronically. It may be something that becomes endemic, that we have to just be careful about" Dr. Anthony Fauci, Nov 12, 2020

The ongoing Covid-19 pandemic has, as of this writing, killed more than 2.5 million people worldwide, including more than 500,000 people in the United States. Social distancing measures aimed at slowing the spread of the virus shut down large sectors of the U.S. economy, leading to the steepest quarterly decline in GDP and the highest level of the unemployment rate since the Great Depression.[1] This economic collapse led many people to question whether the benefits of reduced infections and deaths resulting from social distancing are worth the costs of increased unemployment and decreased output caused by social distancing. This question has spawned an economics literature that analyzes the economic effects of social distancing and weighs the benefits of reduced infection against the costs of reduced economic activity.

Even as Covid vaccines were granted approval in late 2020, leading epidemiologists quoted above warned that the disease is likely to become endemic. In the parlance of economics, an endemic disease is one that persists indefinitely and is a feature of long-run, or steady-state, equilibrium. Most papers in the burgeoning economics literature focus on the evolution of infections and economic activity in the short run. Because these papers are silent about the long run, they are also silent about whether the disease will become endemic. Moreover, these papers typically use versions of the workhorse SIR model that contain neither vital statistics (births and deaths) nor waning immunity, and thus, *by construction*, rule out the possibility of an endemic equilibrium.

To introduce the possibility of an endemic equilibrium in an SIR model, the population

---

[1]The Bureau of Economic Analysis series for quarterly growth rates of real GDP begins in 1947. The decline in 2020Q2 was the steepest quarterly decline on record.

1

of susceptible people (S) must be replenished over time. In this paper, we introduce two channels for the replenishment of S. The first channel operates through the introduction of vital statistics, comprising births, baseline deaths, and excess deaths from the disease. We assume that people are born without any immunity to the disease, so births replenish S. Depending on the specification of the parameters of the model, the steady state can be either endemic (which we label as an endemic equilibrium, EE) or non-endemic (which we label as a disease-free equilibrium, DFE). The population dynamics associated with births and deaths are, admittedly, slow moving so that it takes a long time get close to the steady state. The second channel is the specification that immunity wanes over time. As recovered people (R) lose their immunity, they rejoin S, which can lead to an endemic equilibrium. Again, depending on the specification of parameter values, the steady state can be EE or DFE. Importantly, the waning feature of immunity greatly speeds up the approach to the steady state and thus makes the steady state a relevant object of analysis.

We consider two tools to reduce the spread of the disease. One tool is social distancing, described above, and the other tool is vaccination. In the first several months of the pandemic, before a vaccine was available, social distancing was the only tool available. Therefore, we begin with positive and normative analyses of social distancing. In the normative analysis, the welfare function is the present value of the stream of output minus a penalty that is proportional to the number of excess deaths from the disease. The optimal degree of social distancing mandated by the planner maximizes the welfare function. We also examine a laissez-faire equilibrium that arises when each person decides how much social distancing to practice in the absence of an enforceable mandate.

Our analysis of the pre-vaccine model produces several results. First, the outcome under optimal policy and the laissez-faire outcome both involve a non-trivial amount of social distancing, even in the steady state. Second, in the laissez-faire outcome, the degree of social distancing is smaller, and the amount of infections and excess deaths are higher than under the optimal policy because individuals in the laissez-faire outcome do not internalize the effect of their own infection on the likelihood of an additional infection among susceptible individuals. Third, in the long run, optimal policy leads to an endemic equilibrium (EE)

even though a non-endemic equilibrium (DFE) is attainable. Strikingly, the emergence of an endemic equilibrium holds regardless of the magnitude of the societal cost of an additional death. Indeed, for *any* admissible constellation of model parameters, optimal policy leads to an endemic steady state, as long as the basic reproduction rate of the disease (R_0) in the absence of any social distancing is above one. Fourth, benchmarking the pre-vaccine model to the U.S. experience in the first few months of the pandemic, we find that actual excess deaths were somewhere between the model outcomes under optimal policy on the one hand and laissez-faire on the other hand; in fact, actual experience was closer to the laissez-faire version of the model at the beginning of the pandemic and then moved toward the model under optimal policy, as various states implemented social distancing policies.

We extend the model to include a vaccine that provides immunity to the disease, but the immunity wanes at the same rate that immunity acquired from infection wanes. Optimal policy uses two tools–social distancing and vaccination–to maximize a social welfare function that is the same as described above in the absence of a vaccine, except that the cost of administering the vaccine is subtracted from aggregate output. Again, optimal policy leads to an endemic steady state, though the use of a vaccine changes important features of the steady state. When social distancing is the only tool to limit the spread of the disease, the optimal degree of social distancing, and the consequent reduction in output, can be quite large in the steady state. Adding a vaccine to the policy tool kit relieves social distancing from sole burden of fighting the disease; as a consequence, the optimal degree of social distancing and the associated loss in output become extremely small in the steady state. In addition, the use of a vaccine reduces the rate of excess deaths per year substantially below the rate attained with optimal social distancing alone. In the quantitative implementation of optimal policy for a country with 330 million people, the endemic equilibrium entails about 22,500 deaths per year in the absence of vaccines (a stock of about 187,500 infected people on any given day) but only 65 deaths per year (a stock of 542 infected people on any given day) with the optimal use of vaccines.

Our findings put the potentially alarming comments of Drs. Heymann and Fauci cited above in a new light. While the distinction between endemic equilibria and disease-free

3

equilibria is quite sharp, we demonstrate, perhaps surprisingly, that optimal policy leads to an endemic equilibrium. Therefore, the attainment of a disease-free equilibrium need not be the goal of public health policy. Having said that the optimal steady state will be endemic, we find that in the optimal endemic equilibrium, excess deaths in the United States would be fewer than hundred per year, and the loss of output per person would be essentially undetectable.

The economics literature on the covid-19 pandemic has grown at an explosive rate since the arrival of the coronavirus. We focus on the literature in the tradition of the epidemiological framework based on Kermack et al. (1927), Bailey (1975), Hethcote (2000) andAvery et al. (2020). Three features of the eponymous paradox – that optimal public health policy leads to an endemic steady state – distinguish our paper from the existing literature. First, in contrast to the findings in most of the literature, which depend on numerical computation, the paradox is an analytic result. The advantage of the analytic result is that it holds for an entire set of admissible parameter values and thus does not depend on particular choices of parameter values that are needed in numerical computation.

The second feature of the paradox is that even in the presence of a vaccine that is capable of leading to a steady state with zero infections, the optimal policy will not do so. Of course, to examine this question in the first place, one needs to include vaccines in the tool kit used for public health policy. Most of the recent economics literature, including Berger et al. (2020), Piguillem and Shi (2020), Gonzalez-Eiras and Niepelt (2020b), Gonzalez-Eiras and Niepelt (2020a), Jones et al. (2020), focuses on social distancing as the only tool to slow the spread of the disease. Often when vaccines have been included in the existing literature, the vaccines are treated as an anticipated future tool of public health policy. While forward-looking policy makers would take account of the (stochastic) arrival of a vaccine in deciding on current social distancing mandates, the future arrival of a vaccine is often modeled as an event that immediately transforms all susceptible people into recovered people (Alvarez et al. (2020) , Acemoglu et al. (2020), Eichenbaum et al. (2020)). Indeed, we adopt this assumption in Section 3, when we analyze optimal social distancing in the absence of a vaccine. However, in Section 6, we adopt a more nuanced description of the vaccine in the

evolution of infection dynamics.

The third feature of our analysis, which is needed to even consider the paradox, is that we use an epidemiological model that admits both disease-free and endemic steady states. That is, an endemic steady state resulting from optimal policy is possible, but it is not hardwired. As we have mentioned above, an endemic steady state requires replenishment of the pool of susceptible people either through waning immunity or births of susceptible people. Much of the recent literature includes neither waning immunity nor births, thereby ruling out endemic steady states. Notable papers in this endemic-free framework include Alvarez et al. (2020), Acemoglu et al. (2020), Atkeson (2020), Bethune and Korinek (2020), Cochrane (2020), Fernandez-Villaverde and Jones (2020), Farboodi et al. (2020), and Garibaldi et al. (2020). Cenesiz and Guimarães (2020) and Giannitsarou et al. (2020) include waning immunity in their models, but neither paper includes vaccines and neither paper derives general theoretical results. An older, pre-covid literature examines the possibility of endemic equilibria in the presence of vaccines. Geoffard and Philipson (1997) and Barrett and Hoel (2007) frame the question about the desirability of eradicating a disease. However, Geoffard and Philipson (1997) explicitly ignores steady states that are disease-free.[2] Barrett and Hoel (2007) departs from the SIR framework to examine eradication that occurs in finite time. Goldman and Lightwood (2002) shows that optimal medical treatment of infected people can lead to an endemic steady state in an SIS model, which is a simple special (limiting) case of the SIR model as the hazard rate of waning immunity approaches infinity. Public health policy in Goldman and Lightwood (2002) is confined to the treatment of infected people, whereas the public health policies we examine focus on controlling the flow of new infections, both by decreasing the infection hazard rate through social distancing and by increasing the vaccination rate. Gersovitz and Hammer (2003) also analyze optimal public health policy, excluding vaccination. They remark that optimal policy may lead to a disease-free equilibrium, but rule out such cases in their formal analysis, thereby forcing optimal policy to lead to an endemic steady state.

The planning problem used to derive optimal public health policy takes account of the

---

[2]See their footnote 4: "We ignore the steady state $(S, I, R) = (1, 0, 0)$ with zero prevalence [of the infection]."

externality imposed by infected people on susceptible people. We demonstrate the impact of internalizing this externality by comparing the planner's optimal social distancing (in the absence of a vaccine) to the equilibrium level of social distancing in a "laissez-faire" economy, where the degree of social interaction is chosen freely by individuals, as in the optimization frameworks of Eichenbaum et al. (2020) and Toxvaerd (2020), or informally as in Cochrane (2020).

In Section 1, we present a version of the basic SIR model that incorporates vital demographics, including births, baseline deaths, and excess deaths from the disease; waning immunity; and vaccinations. We analyze the steady state of this model and show which combinations of social distancing and vaccinations lead to disease-free steady states and which combinations of these public health measures lead to endemic equilibria. We specify the impact of social distancing and the infected share of the population on aggregate output per capita in Section 2. In Section 3, we turn from a positive analysis to a normative analysis and derive the steady state that arises from the optimal use of social distancing before the appearance of a vaccine. Then in Section 4, we analyze the steady state that arises in a laissez-faire economy in which individuals choose their own degrees of social distancing in the absence of a vaccine. Section 5 quantitatively implements the short-run equilibria in the early stages of a pandemic under socially optimal social distancing as well as under laissez-faire. In Section 6, we add a vaccine to the tool kit of public health policy and show that even with a vaccine, the optimal steady state is endemic. Concluding remarks are contained in Section 7.

1   SIR Model with Population Growth and Excess Deaths

The total population, a continuum of measure N, is the sum of susceptible people, S, infected people, I, and recovered people, R, who are no longer susceptible to the disease. Let $\phi > 0$ be the birth rate per unit of population per unit of time, $\mu \geq 0$ be the baseline death rate per unit of population per unit of time, $\gamma > 0$ be the recovery rate per infected person per unit of time and $\delta > 0$ be the excess death rate of infected people per unit of time. As in

conventional SIR epidemiological models, the flow of new infections per unit of time is $\beta S \frac{I}{N}$, where $\beta > 0$ is a contagion parameter reflecting the extent of social and business interactions. We augment the SIR model in two important ways: First, we assume that susceptible people are vaccinated with a hazard rate $\lambda \geq 0$. The vaccination transforms susceptible people into recovered people. Second, immunity from the disease, whether acquired by vaccination or by having recovered from the disease, wanes over time with a hazard rate $\chi \geq 0$.

We begin by treating the contagion rate, β, and the vaccination rate, λ, as fixed parameters and conduct a positive analysis of the dynamics and steady state of this augmented SIR model. Beginning in Section 3, we treat β and λ as control variables of public health policy and we analyze the values of these control variables that maximize a planner's objective function.

The differential equations governing the evolution of S, I, and R are

$$\frac{dS}{dt} = \phi N + \chi R - \beta S \frac{I}{N} - \mu S - \lambda S \tag{1}$$

$$\frac{dI}{dt} = \beta S \frac{I}{N} - (\mu + \delta + \gamma) I \tag{2}$$

and

$$\frac{dR}{dt} = \gamma I + \lambda S - \mu R - \chi R. \tag{3}$$

Equation (1) expresses the change in the susceptible population per unit time as the number of births, ϕN, plus the number of recovered people, χR, who lose their immunity and become susceptible, minus the number of people per unit time who leave the population of susceptible people through baseline deaths, μS, becoming infected infected at rate $\beta S \frac{I}{N}$, or getting vaccinated at rate λS. This formulation assumes that all newborns are susceptible. Equation (2) expresses the change in the infected population per unit of time as the number of susceptible people who become infected, $\beta S \frac{I}{N}$, minus the number people who leave the

population of infected people through baseline deaths, μI, excess deaths from the disease, δI, or recovery from the disease, γI. Equation (3) expresses the change in the number of recovered people per unit of time as the sum of people who enter the recovered population either by recovering from the disease, γI, or by vaccination, λS, minus the number of people who leave the recovered population either by baseline death, μR, or by waning immunity, χR.

Throughout our analysis, we assume that

$$\lambda \left(\delta + \gamma - \chi \right) \geq 0. \tag{4}$$

In the absence of a vaccine, $\lambda = 0$, and the restriction in equation (4) has no impact. In the presence of a vaccine, $\lambda > 0$ and the restriction requires $\chi \leq \delta + \gamma$. Note that $(\chi + \mu)^{-1}$ is the expected duration of immunity for an individual since $\chi + \mu$ is the rate at which people lose immunity either through death or by waning of immunity; also, $(\delta + \gamma + \mu)^{-1}$ is the expected duration of infection for an individual, since $\delta + \mu + \gamma$ is the rate at which people leave the infected state, either through death or recovery. Thus, $\chi \leq \delta + \gamma$ is equivalent to requiring that the expected duration of immunity is longer than the expected duration of an infection episode. This restriction is likely to be satisfied. In the quantitative analysis in Section 5, $\delta + \gamma$ is 12.12 and χ is 1, so the restriction in equation (4) is satisfied by a wide margin.

Because population potentially can grow without bound, we work with the population shares $s \equiv \frac{S}{N}$, $i \equiv \frac{I}{N}$, and $r \equiv \frac{R}{N}$, where $s + i + r = 1$. The change in population per unit time is $gN = \phi N - \mu N - \delta I$, which is births, ϕN, less baseline deaths, μN, and less excess deaths, δI, so

$$g = g\left(i\right) \equiv \phi - \mu - \delta i. \tag{5}$$

The change in the susceptible share $s \equiv \frac{S}{N}$ is $\frac{ds}{dt} = \frac{1}{N}\frac{dS}{dt} - gs$, so

$$\frac{ds}{dt} = \phi + \chi r - \beta s i - \left(\mu + g + \lambda \right) s, \tag{6}$$

Similarly,

$$\frac{di}{dt} = [\beta s - (\delta + \gamma + \mu + g)]\, i \tag{7}$$

and

$$\frac{dr}{dt} = \gamma i + \lambda s - (\mu + g + \chi)\, r. \tag{8}$$

Define the *effective* reproduction rate, $\mathcal{R}_t$, as the ratio of βsi, which is the gross increase in the infection share i shown by the first term on the right hand side of equation (7), to $(\delta + \gamma + \mu + g)\, i$, which is the gross decrease in share i resulting from disease-caused deaths (δi), baseline deaths (μi), the recovery of infected people (γi) and population growth (gi). Therefore,

$$\mathcal{R}_t \equiv \frac{\beta s}{\delta + \gamma + g + \mu} \tag{9}$$

and equation (7) can be rewritten as

$$\frac{di}{dt} = (\delta + \gamma + g + \mu)\,(\mathcal{R}_t - 1)\, i, \tag{10}$$

where $\delta + \gamma + g + \mu = \delta + \gamma + \phi - \delta i > 0$. Equation (10) shows that the sign of $\mathcal{R}_t - 1$ determines whether a positive infection share, $i > 0$, is increasing or decreasing.

Define the *basic* reproduction rate $\mathcal{R}_0$ as the value of the effective reproduction rate when $s = 1$, as would be the case at the beginning of an epidemic if no one has any immunity from the disease before the epidemic. Therefore,

$$\mathcal{R}_0 \equiv \frac{\beta}{\delta + \gamma + g + \mu}. \tag{11}$$

The basic reproduction rate, $\mathcal{R}_0$, plays a pivotal role in determining whether the long-run equilibrium is a disease-free equilibrium (DFE) with a zero measure of infected people or an endemic equilibrium (EE) with a positive measure of infected people. Note that since the

9

growth rate, g, is endogenous, both $\mathcal{R}_0$ and $\mathcal{R}_t$ are endogenous.

Remarkably, the expression for $\frac{di}{dt}$ in equation (7) does not directly depend on either the vaccination parameter, λ, or the waning immunity parameter, χ. Since the inflows to and outflows from the pool of infected people are unaffected by λ and χ, the effective reproduction rate, $\mathcal{R}_t$, and the basic reproduction rate, $\mathcal{R}_0$, do not directly depend on either λ or χ. The expression for the growth rate of the economy in equation (5), $g = g(i) \equiv \phi - \mu - \delta i$, also does not depend directly on λ or χ, though the endogenous value of the growth rate, and consequently the values of the effective reproduction rate and the basic reproduction rate, are affected by λ and χ to the extent that λ and χ affect the infection rate i.

1.1 Steady-State Equilibria

In a steady state, S, I, R, and N all grow at the rate g, so s, i, and r are constant and equal s^*, i^*, and r^*, respectively. Throughout, an asterisk (*) denotes steady-state values of variables.

We begin the analysis of the steady state by deriving an upper bound on i^*. Set $\frac{dr}{dt}$ in equation (8) equal to zero to obtain

$$\gamma i^* = (\mu + g + \chi) r^* - \lambda s^* \leq (\mu + g + \chi) r^* \leq (\phi + \chi) r^*, \tag{12}$$

where the final inequality follows from the definition of the growth rate $g(i)$ in equation (5). Equation (12) implies that $r^* \geq \frac{\gamma}{\phi+\chi} i^*$, so that $1 \geq i^* + r^* \geq i^* + \frac{\gamma}{\phi+\chi} i^* = \frac{\phi+\chi+\gamma}{\phi+\chi} i^*$, which proves the following proposition.

Proposition 1 $0 \leq i^* < \frac{\phi+\chi}{\gamma+\phi+\chi} < 1.$

In addition to providing a quantitative upper bound on i^*, Proposition 1 emphasizes the importance of $\phi + \chi > 0$ for the existence of an endemic steady state. An endemic steady state requires replenishment of the pool of susceptible people either through births ($\phi > 0$) or through waning immunity ($\chi > 0$). Without births or waning immunity – strictly speaking, in the limit as $\phi + \chi$ approaches zero, the upper bound in Proposition 1 implies that i^* approaches zero – the steady state is not endemic.

10

To facilitate the analysis of the steady state, we define a critical value of the contagion parameter β

$$\beta_c \equiv \delta + \gamma + \phi. \tag{13}$$

As will become evident from Proposition 4, in the absence of a vaccine, that is, when $\lambda = 0$, the steady state is DFE if $\beta \leq \beta_c$ and is EE if $\beta > \beta_c$. In the presence of a vaccine, the value of β that is the boundary between DFE and EE states is $\widetilde{\beta}_c(\lambda)$, which is defined as

$$\widetilde{\beta}_c(\lambda) \equiv \left(1 + \frac{\lambda}{\phi + \chi}\right)\beta_c \equiv \left(1 + \frac{\lambda}{\phi + \chi}\right)(\delta + \gamma + \phi). \tag{14}$$

It is also useful to define

$$\eta \equiv g + \mu + \chi = \phi + \chi - \delta i \leq \phi + \chi, \tag{15}$$

where the final equality uses equation (5) to replace $g + \mu$ by $\phi - \delta i$. The variable η captures the forces that tend to decrease the recovered share, r, over time. Specifically, r decreases because of baseline deaths at rate μ, waning immunity at rate χ, and dilution due to population growth at rate g.

In the steady state, $\frac{ds}{dt} = \frac{di}{dt} = \frac{dr}{dt} = 0$. Since $s + i + r \equiv 1$, it is sufficient to characterize the conditions for any two of these three time derivatives to be zero. Set $\frac{di}{dt}$ in equation (7) equal to zero, use the definition of the critical value of β from equation (13) and use equation (5) to replace $g + \mu$ by $\phi - \delta i$ to obtain

$$[\beta s^* - (\beta_c - \delta i^*)]\, i^* = 0. \tag{16}$$

Set $\frac{dr}{dt}$ in equation (8) equal to zero and use the definition of η in equation (15) to obtain

$$\gamma i^* + \lambda s^* = \eta^* r^*. \tag{17}$$

11

Equation (17) implies[3]

$$\eta^* > 0. \tag{18}$$

Equations (16) and (17), together with $s^* + i^* + r^* = 1$, lead directly to the following two propositions.

Proposition 2 *If $i^* = 0$, then*

1. $s^* = 1 - \frac{\lambda}{\phi + \chi + \lambda} > 0$

2. $r^* = \frac{\lambda}{\phi + \chi + \lambda} < 1.$

Proposition 3 *If $i^* > 0$, then*

1. $s^* = \frac{\beta_c - \delta i^*}{\beta} = \mathcal{R}_0^{*-1}$

2. $i^* = \frac{\eta^*}{\eta^* + \gamma}\left(1 - \mathcal{R}_0^{*-1}\right) - \frac{\lambda}{\eta^* + \gamma}\mathcal{R}_0^{*-1}$

3. $r^* = \frac{\gamma}{\eta^* + \gamma}\left(1 - \mathcal{R}_0^{*-1}\right) + \frac{\lambda}{\eta^* + \gamma}\mathcal{R}_0^{*-1}$

The expressions for i^* and r^* in Proposition 3 involve the endogenous variable η^*. Lemma 4 in the appendix proves that if $i^* > 0$, then η^* is the unique positive root of the following quadratic equation

$$q\left(\eta\right) \equiv \left[\left(\beta - \delta\right)\eta + \beta\gamma - \delta\lambda\right]\left(\eta - \left(\phi + \chi\right)\right) + \delta\left[\beta\eta - \left(\eta + \lambda\right)\beta_c\right]. \tag{19}$$

Proposition 4 *If $\beta \leq \widetilde{\beta}_c(\lambda) \equiv \left(1 + \frac{\lambda}{\phi + \chi}\right)\beta_c$, then $i^* = 0$, s^* and r^* are as stated in Proposition 2, and $\mathcal{R}_0^* \leq 1 + \frac{\lambda}{\phi + \chi}$. If $\beta > \widetilde{\beta}_c(\lambda) \equiv \left(1 + \frac{\lambda}{\phi + \chi}\right)\beta_c$, then $i^* > 0$, s^*, i^*, and r^* are as stated in Proposition 3, and $\mathcal{R}_0^* > 1 + \frac{\lambda}{\phi + \chi}$.*

Define

$$\lambda_h\left(\beta^*\right) \equiv \min\left\{\lambda \geq 0 : i^*\left(\beta^*\right) = 0\right\} \tag{20}$$

---

[3]If $\gamma i^* + \lambda s^* > 0$, then $\eta^* r^* > 0$, which implies $\eta^* > 0$. If $\gamma i^* + \lambda s^* = 0$, then $i^* = 0$, so equation (15) implies $\eta^* = \phi + \chi > 0$.

12
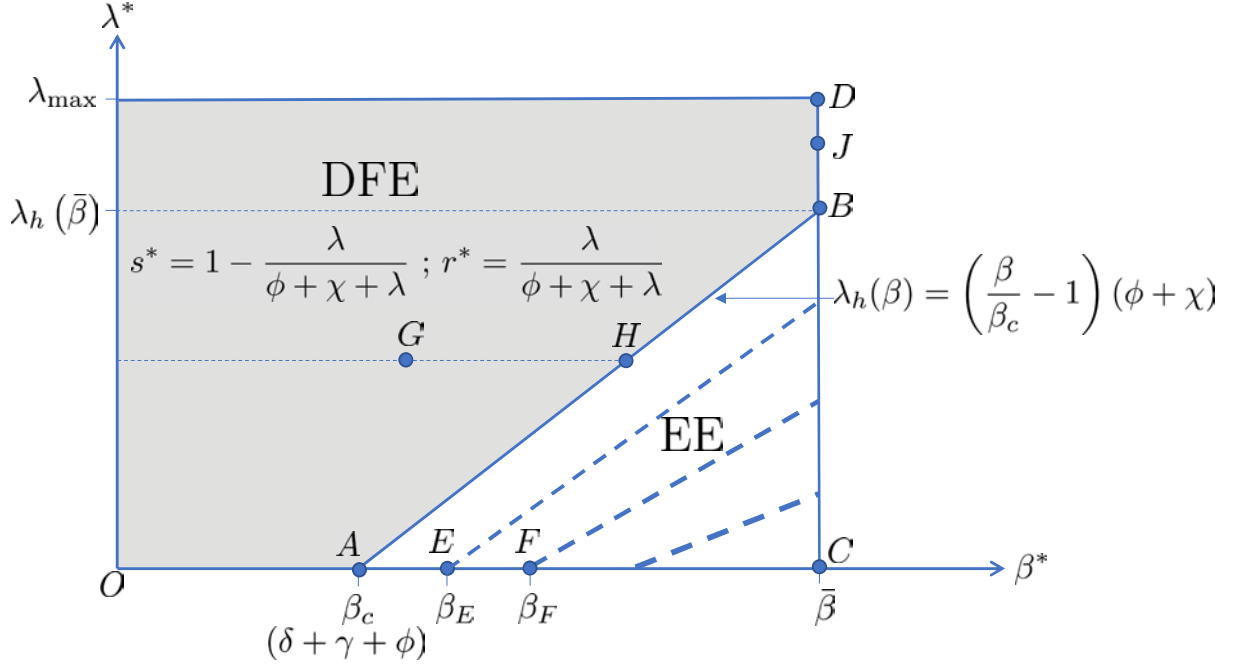
Figure 1: An illustration of disease-free and endemic equilibria in the presence of vaccination.

as the minimum vaccination rate λ that achieves a DFE steady state when the contagion parameter in the steady state is β^*.

Proposition 5 *If $\beta^* \geq \beta_c$, then*

1. $\lambda_h\left(\beta^*\right) = \left(\frac{\beta^*}{\beta_c} - 1\right)\left(\phi + \chi\right)$

2. *if $\lambda^* = \lambda_h\left(\beta^*\right)$, then $s^* = \frac{\beta_c}{\beta^*}$, $i^* = 0$ and $r^* = 1 - \frac{\beta_c}{\beta^*}$.*

The following proposition shows how a positive steady-state infection rate depends on the steady-state contagion parameter β^* and the steady-state vaccination rate λ^*.

Proposition 6 *If $\beta^* > \widetilde{\beta}_c\left(\lambda^*\right) \equiv \left(1 + \frac{\lambda^*}{\phi + \chi}\right)\beta_c$, then*

1. $\frac{di^*}{d\beta^*} = \beta^{*-1}\left(\eta^* + \lambda^*\right)\left[\left(1 - i^*\right)\delta + \gamma + \phi\right]\left[q'\left(\eta^*\right)\right]^{-1} > 0$

2. $\frac{di^*}{d\lambda^*} = -\left[\left(1 - i^*\right)\delta + \gamma + \phi\right]\left[q'\left(\eta^*\right)\right]^{-1} < 0$.

13

Figure 1 illustrates the results in the preceding propositions and corollaries for combinations of the steady-state contagion parameter β^* on the horizontal axis and the steady-state vaccination parameter λ^* on the vertical axis. The upward-sloping line segment $\overline{AB}$ is a graph of the function $\lambda_h(\beta)$ in Proposition 5. It is the boundary between (β^*, λ^*) pairs that lead to DFE steady states (in the shaded region above and to the left of $\overline{AB}$) and (β^*, λ^*) pairs that lead to EE steady states (below and to the right of $\overline{AB}$). Along $\overline{AB}$, the steady states are DFE. Thus, points with sufficiently high vaccination parameter λ^* or sufficiently low contagion parameter β^* will have DFE steady states. However, if λ^* is sufficiently low and β^* is sufficiently high, so that the (β^*, λ^*) pair is below $\overline{AB}$, then the steady state is EE.

Figure 1 shows that if $\beta^* \leq \beta_c \equiv \delta + \gamma + \phi$, the steady state is DFE for any admissible value of λ^*, including zero; thus, a DFE steady state can be attained without a vaccine. Figure 1 also shows that if $\lambda^* \geq \lambda_h(\overline{\beta})$, the steady state is DFE even when $\beta^* = \overline{\beta}$ so a DFE steady state can be attained without social distancing.

Within the DFE region, $i^* = 0$ and the values of s^* and r^* depend linearly on λ^* but are invariant to β^* (Proposition 2). Within this region, a decrease in λ^* reduces the rate at which people transfer directly from the susceptible status to the recovered status, and this reduced rate is reflected in an increased value of s^* and a decreased value of r^*. Indeed, if the vaccination parameter is reduced to $\lambda^* = 0$ when $\beta^* \leq \beta_c$, then the steady state is characterized by $s^* = 1$ and $i^* = r^* = 0$.

Within the EE region in Figure 1, the steady-state infection rate, i^*, is positive. Proposition 6, which states that i^* is increasing in β^* and decreasing in λ^*, suggests that increasing β^* and increasing λ^* by appropriate amounts will leave i^* unchanged. That is, there is an upward-sloping iso-infection locus of (β^*, λ^*) pairs for any $i^* > 0$. As shown in Appendix A.1.1., the iso-infection loci are, in fact, linear. The upward-sloping dashed lines in the EE region of Figure 1 are iso-infection lines.[4] Moving up and to the right along a given iso-infection line, s^* decreases and r^* increases, while, of course, i^* remains unchanged.[5]

---

[4]Appendix A.1.1. shows that the iso-infection line through, say, point E on the horizontal axis, where $\beta^* = \beta_E$ and $\lambda^* = 0$, is $\lambda^* = \frac{\eta_E^*}{\beta_E}\beta^* - \eta_E^*$, where η_E^* is the positive root of $q(\eta) = 0$ when $\beta^* = \beta_E$ and $\lambda^* = 0$ in the quadratic function $q(\eta^*)$ in equation (19).

[5]In the EE region, Statement 1 of Proposition 3 is that $s^* = \mathcal{R}_0^{*-1}$, which implies $s^* = \frac{\delta + \gamma + g(i^*) + \mu}{\beta^*}$. Since i^* is constant along an iso-infection line, $g(i^*)$ is also constant, and hence s^* is inversely proportional

14

As one moves to the right in the EE region, the iso-infection lines become less steep and, as indicated by the increasing thickness of the iso-infection lines, the value of i^* increases, reaching a maximum value at Point C.

2 Effect of Social Distancing on Output

We will use the term social distancing to denote non-medical public health policies aimed at reducing the contagion parameter β. Social distancing policies can range from making hand sanitizer readily available in public places and wearing of face coverings to limiting or outright closing various social and productive activities where people gather, such as schools, offices, factories, personal care services, sports and entertainment venues, and restaurants and bars. We will specify aggregate output per capita to be an increasing function of β. To the extent that social distancing reduces β, output is a decreasing function of social distancing.

Define z as an index of a person's social and productive interactions. The amount of output produced by that person per unit of time is strictly increasing in z for $z < \overline{z}$ and is constant for $z \geq \overline{z}$. The contagion parameter β is strictly increasing in z, so we write the amount of output per unit of time produced by that person as $y(\beta)$ with $y'(\beta) > 0$ for $\beta < \overline{\beta}$ and $y'(\beta) \equiv 0$ for $\beta \geq \overline{\beta}$, where $\overline{\beta}$ is the value of β when $z = \overline{z}$.

We assume that infected people do not produce output so that aggregate output per capita is

$$Y = sy(\beta_s) + ry(\beta_r) \tag{21}$$

where β_s and β_r are the contagion parameters for susceptible and recovered people, respectively. If $\chi > 0$, so that recovered people eventually become susceptible, we need to consider whether susceptible and recovered people can be distinguished from each other. We analyze two extreme cases. If susceptible and recovered people can be costlessly distinguished from each other, then there are distinct values of the contagion parameter, β_s and β_r, respectively,

---

to β^* along an iso-infection line.

for these two groups. In this case, there is no need for recovered people to restrain their interactions below $\overline{z}$, so $\beta_r = \overline{\beta}$. Alternatively, if the cost of distinguishing susceptible people from recovered people is sufficiently high, then $\beta_r = \beta_s$. We can capture both cases with the specification $\beta_r \equiv \zeta\beta_s + (1-\zeta)\overline{\beta}$, $\zeta \in \{0,1\}$, where $\zeta = 1$ if susceptible people and recovered people are indistinguishable from each other and $\zeta = 0$ if susceptible and recovered people can be costlessly distinguished from each other. Henceforth, we use β to represent β_s, which implies

$$\beta_r \equiv \zeta\beta + (1-\zeta)\overline{\beta}, \quad \zeta \in \{0,1\}. \tag{22}$$

Equations (21) and (22) imply that aggregate output per capita is

$$Y(\beta, s, r) = sy(\beta) + ry(\beta_r) = sy(\beta) + ry\left(\zeta\beta + (1-\zeta)\overline{\beta}\right), \quad \zeta \in \{0,1\}. \tag{23}$$

Equation (23) implies the following partial derivatives

$$Y_\beta(\beta, s, r) = (s + \zeta r)y'(\beta), \quad \zeta \in \{0,1\} \tag{24}$$

and

$$Y_s(\beta, s, r) = y(\beta) \leq y\left(\zeta\beta + (1-\zeta)\overline{\beta}\right) = Y_r(\beta, s, r), \tag{25}$$

where the weak inequality holds as an equality if $\zeta = 1$ or if $\beta = \overline{\beta}$.

3   Optimal Social Distancing without a Vaccine

In the beginning stages of a pandemic, before effective treatments or vaccines are available, a forward-looking social planner anticipates that medical and scientific developments will bring relief from the pandemic. To incorporate the anticipation of future relief, without endowing the planner with too much knowledge of when or in what form that relief will arrive, we follow Alvarez et al. (2020) and model the planner as assuming that there will be a "breakthrough

date," denoted T, when medical therapies lead to complete and instantaneous recovery of all infected people and a vaccine prevents any new infections, so that the disease is completely eradicated. To capture the uncertainty about the timing of the breakthrough date, we let p be the constant hazard rate of arrival of the breakthrough. This formulation of the breakthrough is a simplification. We use the simplification merely to guide the dynamics of policy in the pre-vaccine phase of the pandemic. In Section 6, we analyze the cum-vaccine phase of the pandemic and adopt a more realistic specification of the role of the vaccine in taming the spread of the disease. In particular, we specify that the immunity provided by the vaccine, as well as immunity obtained by recovery from the disease, wanes over time, replenishing the population of susceptible people.

From T onward, $i = 0$, $g\left(i\right) = \phi - \mu$, and optimal $\beta = \overline{\beta}$. Therefore,

$$N_T \widetilde{V} \equiv N_T \int_T^\infty y\left(\overline{\beta}\right) e^{-(\rho - (\phi - \mu))(t-T)} dt = N_T \frac{y\left(\overline{\beta}\right)}{\rho - (\phi - \mu)} \tag{26}$$

is the present value, discounted at rate $\rho > 0$, of aggregate output from date T onward. To ensure that $\widetilde{V}$ is finite, assume that

$$\rho > \phi - \mu. \tag{27}$$

Before T, the optimal time path of β balances the benefit of reducing β in terms of reducing excess deaths against the cost of reducing β in terms of lost output. This tradeoff is reflected in the objective function

$$\max_{\beta_u, t \leq u \leq T} E_t \left\{ \int_t^T N_u e^{-\rho u} \left[Y\left(\beta_u, s_u, r_u\right) - \omega \delta i_u \right] du + N_T e^{-\rho(T-t)} \widetilde{V} \right\}, \tag{28}$$

where $0 < \omega < \infty$ is the weight the planner places on an excess death relative to a unit of aggregate output per capita.

Using $N_u = N_t \exp\left(\int_t^u g_z dz\right)$, the objective function in (28), per unit of population, N_t,

at time t is

$$V\left(s_t, i_t, r_t\right) = \max_{\beta_u, t \leq u \leq T} E_t \left\{ \int_t^T e^{-\int_t^u (\rho - g_z) dz} \left[Y\left(\beta_u, s_u, r_u\right) - \omega \delta i_u \right] du + \widetilde{V} e^{-\int_t^T (\rho - g_z) dz} \right\}. \quad (29)$$

Recognizing that T is a random value with constant hazard p, the Bellman equation associated with the objective function in equation (29) is

$$\left(\rho - g\left(i\right)\right) V\left(s, i, r\right) = \max_{\beta} \left\{ \begin{array}{c} Y\left(\beta, s, r\right) - \omega \delta i \\ +V_s \frac{ds}{dt} + V_i \frac{di}{dt} + V_r \frac{dr}{dt} + p\left(\widetilde{V} - V\left(s, i, r\right)\right) \end{array} \right\}. \quad (30)$$

The left hand side of (30) is the required return per unit time, which is the growth-adjusted discount rate, $\rho - g\left(i\right)$, multiplied by V. The right hand side of this equation is the expected return, which comprises the instantaneous flow of welfare, $Y\left(\beta, s, r\right) - \omega \delta i$, and the expected change in $V\left(i, s, r\right)$, which consists of the change resulting from changes in the state variables, s, i, and r, $V_s \frac{ds}{dt} + V_i \frac{di}{dt} + V_r \frac{dr}{dt}$, and the expected change associated with the breakthrough, $p\left(\widetilde{V} - V\left(s, i, r\right)\right)$.

To obtain expressions for the dynamic behavior of $V_j(s, i, r)$, $j \in \{s, i, r\}$, differentiate both sides of (30) with respect to j and use $\frac{dV_j}{dt} = V_{js} \frac{ds}{dt} + V_{ji} \frac{di}{dt} + V_{jr} \frac{dr}{dt}$. Appendix A shows that these calculations lead to

$$\left(\rho + \mu + p\right) V_s = y\left(\beta\right) + \frac{dV_s}{dt} + \beta i \left(V_i - V_s\right) \quad (31)$$

$$\left(\rho + \mu + p + \delta\right) V_i = -\omega \delta + \frac{dV_i}{dt} + \gamma \left(V_r - V_i\right) - \delta \left[V - \left(s V_s + i V_i + r V_r\right)\right] - \left(V_s - V_i\right) \beta s \quad (32)$$

and

$$\left(\rho + \mu + p\right) V_r = y\left(\beta_r\right) + \frac{dV_r}{dt} + \chi \left(V_s - V_r\right), \quad (33)$$

where β_r is defined in equation (22).

18

In (31), the effective discount rate on the left side is $\rho + \mu + p$. As in models of uncertain lifetimes going back to Yaari (1965), the effective discount rate includes the rate of pure time preference, ρ, and the instantaneous hazard rate of death. Here, the hazard rate of death is the baseline death rate, μ, plus the hazard rate, p, that T will arrive, terminating the regime in which the disease is present. Thus, the left hand side of (31) is the required return associated with increasing s by one unit. It is equated with the expected return on the right hand side, which consists of the output $y(\beta)$ produced by an additional unit of s and the change in valuation reflecting: (1) the passage of time, $\frac{dV_s}{dt}$, and (2) the increased hazard of becoming infected, βi, multiplied by the change in valuation, $V_i - V_s$, as a person moves from susceptible to infected. The interpretation of (33) is similar, except that $\chi(V_s - V_r)$ is the impact of a unit increase in r on the transition from recovered status to susceptible status as recovered people lose immunity at rate χ, multiplied by $V_s - V_r$, which is the impact on welfare of such a transition.

The interpretation of (32) is more complicated. On the left hand side, the effective discount rate, $\rho + \mu + p + \delta$, includes δ because a unit increase in i increases excess deaths by δ. The first three terms on the right hand side are similar to the terms on the right hand side of (31): The first term reflects that a unit increase in i increases deaths by δ, which reduces the flow of welfare by $\omega\delta$; the second term, $\frac{dV_i}{dt}$, captures the change in V_i with the passage of time; and the third term, $\gamma(V_r - V_i)$, is the hazard rate γ of switching from status i to status r, multiplied by the change in valuation, $V_r - V_i$, associated with that change. The fourth term reflects that a unit increase in i reduces the population change by δN, reducing the aggregate flow of utility by $\delta\frac{\partial}{\partial N}NV\left(\frac{S}{N}, \frac{I}{N}, \frac{R}{N}\right)$, which equals[6] $\delta\left[V - (sV_s + iV_i + rV_r)\right]$. The fifth term, $-(V_s - V_i)\beta s$, reflects an important externality, namely, that an increase in the infection share i increases by βs the hazard rate that a susceptible person will become infected. The planner takes account of this externality by including the change in welfare associated with this new infection $-(V_s - V_i)$ multiplied by βs.

---

[6] $\frac{\partial}{\partial N}NV\left(\frac{S}{N}, \frac{I}{N}, \frac{R}{N}\right) = V(s, i, r) - sV_s(s, i, r) - iV_i(s, i, r) - rV_r(s, i, r)$.

The first-order condition for optimal β is

$$(s + \zeta r) \, y' \, (\beta) = (V_s - V_i) \, si. \tag{34}$$

The optimal value of β equates the marginal benefit and marginal cost increasing β. The marginal benefit is the increase in output per capita, $Y_\beta \, (\beta, s, r) = (s + \zeta r) \, y' \, (\beta)$. Since a unit increase in β increases the infection rate by si, which reduces s by si units and increases i by si units, the marginal cost of increasing β is $(V_s - V_i) \, si$.

3.1   Steady State Under Optimal Policy

To better characterize optimal social distancing policy, captured by β, prior to the arrival of the vaccine, we first consider the situation in which $\frac{ds}{dt}, \frac{di}{dt}, \frac{dr}{dt}, \frac{dV_s}{dt}, \frac{dV_i}{dt}$, and $\frac{dV_r}{dt}$ are all equal to zero. We refer to this situation as a steady state. [7] As we show in the quantitative analysis in Section 5, the transition to this pre-vaccine steady state is rapid, so the conclusions presented in this section are applicable after the first few months of the pandemic.

Let β^* denote the steady-state value of β and let β^*_{OP} denote the value of β^* under the optimal policy.

Lemma 1 *Assume that no vaccine is yet available. If $\overline{\beta} > \beta_c \equiv \delta + \gamma + \phi$, then $V_s^* - V_i^*$ is positive and finite.*

Lemma 1 helps prove the following proposition.

Proposition 7 *Assume that no vaccine is yet available. If $\overline{\beta} > \beta_c \equiv \delta + \gamma + \phi$, and $\omega > 0$ is finite, then $\beta_c < \beta^*_{OP} < \overline{\beta}$. Therefore, under the optimal policy, the steady state is EE.*

Proof of Proposition 7. We prove this proposition by contradiction. We show that both $\beta^*_{OP} \leq \beta_c$ and $\beta^*_{OP} \geq \overline{\beta}$ are impossible. **Case I**: $(\beta^*_{OP} \nleq \beta_c)$: For any $\beta^* \leq \beta_c$, $i^* = 0$ and therefore Proposition 2 implies $(s^* + \zeta r^*) \, y' \, (\beta^*) > 0$. Therefore, since $V_s^* - V_i^* > 0$ is finite

---

[7]This analysis (unlike the steady-state analysis in Section 6) is confined to the period of time before the breakthrough date, T, and hence is a "pseudo steady state."

(Lemma 1), the marginal cost of β, $(V_s^* - V_i^*) s^* i^*$, equals zero and the marginal benefit of β, $(s^* + \zeta r^*) y'(\beta) > 0$, exceeds the marginal cost. Therefore, $\beta_{OP}^* > \beta_c$. **Case II**: $(\beta_{OP}^* \not\geq \overline{\beta})$: Since $V_s^* - V_i^* > 0$ and, from Proposition 4, $i^* > 0$, the marginal cost of β, $(V_s^* - V_i^*) s^* i^*$, is positive and exceeds the marginal benefit of β, $(s^* + \zeta r^*) y'(\beta^*) = (s^* + \zeta r^*) y'(\overline{\beta}) = 0$. Therefore, $\beta_{OP}^* < \overline{\beta}$. ∎

Proposition 7 implies that under the optimal policy, the steady-state value of i^* in the absence of a vaccine is positive and thus the equilibrium is endemic. Nevertheless, the proposition also states that $\beta_{OP}^* < \overline{\beta}$, so that the optimal policy never stops imposing some degree of social distancing (at least, prior to the arrival of a vaccine).

4 Laissez Faire in the Pre-Vaccine Regime

In the absence of centralized policy to control β, individual susceptible people may choose to limit their interactions to reduce their own risks of becoming infected. In this section, we analyze the optimal choice of β for individuals in each of the three states of health. To focus the analysis, we assume that $\zeta = 0$, so that individuals know whether they are susceptible or recovered. In this case, output per susceptible person is $y(\beta)$ and equation (22) implies that output per recovered person is $y(\overline{\beta})$.

Consider the decision of a susceptible person, who knows that in the future she may become infected and subsequently may recover from the disease. Using backward induction, first consider the recovered stage of life. The expected present value, discounted at rate ρ, of a recovered person's earnings until her death, or the arrival of T, whichever comes first, is v^R and satisfies

$$(\rho + \mu + p) v^R = y(\overline{\beta}) + \frac{dv^R}{dt} + \chi (v^S - v^R). \tag{35}$$

For an infected person, the effective discount rate, $\rho + \mu + \delta + p$, includes the hazard rate, δ, that the person dies from the disease. Let v^I be the value of being in the infected state. The instantaneous flow of welfare, $-\omega\delta$, and the hazard-weighted change in value

when recovering from the disease, $\gamma \left(v^R - v^I \right)$, satisfy

$$(\rho + \mu + p + \delta) \, v^I = -\omega \delta + \frac{dv^I}{dt} + \gamma \left(v^R - v^I \right). \tag{36}$$

A susceptible person chooses how much to expose herself to infection according to the Bellman equation

$$(\rho + \mu + p) \, v^S = \max_{\beta} \left\{ y \left(\beta \right) + \frac{dv^S}{dt} + \beta i \left(v^I - v^S \right) \right\}. \tag{37}$$

The first-order condition for the maximization in (37) is

$$y' \left(\beta \right) = \left(v^S - v^I \right) i. \tag{38}$$

Now compare an individual's v^R, v^I, and v^S with the derivatives of the planner's value function V_r, V_i, and V_s, respectively. Comparing (35) with (33) and observing that $\zeta = 0$ implies that $\beta_r = \overline{\beta}$ implies that $v^R = V_r$. Also, (37) has the same form as (31) and the first-order condition (38) has the same form as (34) with $\zeta = 0$, where V_s corresponds to v^S and V_i corresponds to v^I.

The optimal values of β in the two problems differ because the expression for v^I in (36) has a different form than the expression for V_i in (32). The last two terms on the right hand side of (32) have no counterpart in (36). Because v^I differs from V_i, the laissez-faire choice of β differs from the planner's choice, so social welfare can be improved by mandating β rather than relying on individual precaution.

The two terms on the right side of (32) that are responsible for the discrepancy between V_i and v^I are $-\delta \left[V - (sV_s + iV_i + rV_r) \right]$, and $-\left(V_s - V_i \right) \beta s$. The first term is related to the fact that the planner takes into account the impact of i on population growth. The second term, $\left(V_s - V_i \right) \beta s$, reflects an important externality, as discussed earlier. An infected individual does not internalize the contagion of her infection, while the planner does. This externality is the fundamental reason that public health policy is useful.

Proposition 8 *In the absence of a vaccine, there exists a unique steady-state laissez-faire*

equilibrium, $\beta^*_{LF} < \overline{\beta}$.

Proposition 8 shows that under laissez faire, and in the absence of a vaccine, individuals choose to reduce their social and productive interactions relative to the pre-pandemic world. However, as shown in the following proposition, they do not reduce their interactions as much as in the socially optimal steady state in Section 3.

Proposition 9 *In the absence of a vaccine, in a non-negative neighborhood of $\delta = 0$, $\beta^*_{LF} > \beta^*_{OP}$.*

5   Quantitative Analysis of the Early Stage of the Coronavirus Pandemic

Before we analyze optimal policy in the presence of a vaccine, we illustrate the short-run quantitative behavior of the model in the absence of a vaccine ($\lambda \equiv 0$) at the beginning of the coronavirus pandemic. The goal of this section is two-fold. First, we compare the behavior of the model to the actual evolution of the pandemic during its early stage. Second, we demonstrate our earlier claim that the model approaches its steady state quickly.

We set $\phi = 0.015$, which is the sum of the annual birth and net immigration rates, and $\mu = 0.01$, which is the annual death rate in the United States.[8] The values of disease-related parameters are based on the US experience with Covid-19. We set $\gamma = 12$ to reflect that the average person who recovers was infected for about one month. We set $\delta = 0.01\gamma$ $= 0.12$ to reflect that infected people are about 1% as likely to die from the disease as to recover from it. Therefore, the critical value $\beta_c \equiv \delta + \gamma + \phi = 12.135$. To calibrate $\overline{\beta}$, we use (11), the definition of β_c, and the fact that δi is so much smaller than β_c to obtain $\overline{\beta} = (\beta_c - \delta i)\overline{\mathcal{R}}_0 \approx \beta_c \overline{\mathcal{R}}_0$, where $\overline{\mathcal{R}}_0$ is the maximal value of $\mathcal{R}_0$ observed at the beginning of the pandemic before any individual actions or any public health policies to reduce contagion. We use the high end of estimates for $\mathcal{R}_0$ across US states in the last week of February 2020

---

[8]Source: United Nations Population Division, (2015-2020).

and set $\overline{\mathcal{R}}_0 = 3.5$, which implies $\overline{\beta} = 42.473$.[9] The value of the discount rate used by the planner and by individuals, ρ, is set to 0.03. Consistent with the analysis of laissez-faire in Section 4, $\zeta = 0$.

We specify the production function $y(\beta)$ to be quadratic with maximal value $y(\overline{\beta})$ normalized to one. Specifically, for $\beta \leq \overline{\beta}$, $y(\beta) = 1 - \alpha(\overline{\beta} - \beta)^2$, $y'(\beta) = 2\alpha(\overline{\beta} - \beta)$, and the "output gap" is $y(\overline{\beta}) - y(\beta) = \alpha(\overline{\beta} - \beta)^2$.[10] Let $\Delta \equiv \frac{y(\overline{\beta}) - y(\beta_c)}{y(\overline{\beta})} = \alpha(\overline{\beta} - \beta)^2$ be the reduction, as a fraction of $y(\overline{\beta})$, in y when β is reduced from its pre-pandemic level, $\overline{\beta}$, to β_c, the level of β at which $\mathcal{R}_0 = 1$, which was approximately the value of the reproduction rate for most states during the second quarter of 2020. We set $\Delta = 0.09$ to match the 9% drop in output in that quarter. Therefore, $\alpha(\overline{\beta} - \beta_c)^2 = \Delta$, which implies $\alpha = \frac{\Delta}{(\overline{\beta} - \beta_c)^2} = 9.779 \times 10^{-5}$. At the time of the writing of this paper, it is not known how long the immunity of the recovered patients lasts. In our baseline case, $\chi = 1$, which implies that immunity has an average duration of one year. To illustrate the quantitative role of χ, in Tables 1 and 2 we also present results for $\chi = 0.5$ and $\chi = 0$.

To calibrate ω we use the concept of "Quality Adjusted Life Year" (QALY), defined as the value of extending quality life by an extra year. The World Health Organization consensus is that QALY is 1 - 3 times annual GDP per capita.[11] Using a discount rate of $\rho + \mu = 0.03 + 0.01$ to discount the foregone stream of 1 QALY per year over the lost years of life implies that the present value of the losses from an excess death is 25 QALY. Assigning a value of 2 times GDP per capita to each QALY implies $\omega = 50$ times GDP per capita. With $y(\overline{\beta}) = 1$, we set ω equal to 50.[12]

Finally, i_0, the infection share of the population on the initial day of our simulation (March 1), is chosen so that the daily excess death count implied by the model matches the 270 daily Covid-related deaths observed three weeks later (March 22). The daily death rate

---

[9]Source: Estimates of the reproduction rate provided by the websites http://rt.live and http://epiforecasts.io. For instance, rt.live estimates that on February 27, 2020, the effective reproduction rate of Covid 19 was 3.98 in New Jersey and approximately 3.6 in New York and Illinois.

[10]For $\beta \geq \overline{\beta}$, $y(\beta) \equiv 1$.

[11]Source: "Overview of the ICER value assessment framework and update for 2017-2019", p. 15, available at: https://icer.org/wp-content/uploads/2020/10/ICER-value-assessment-framework-Updated-050818.pdf

[12]Alvarez et al. (2020), citing Hall et al. (2020), set $\omega = 20$, but note that this value "is on the low range of the estimates in the literature." Our fundamental result that optimal policy leads to an endemic equilibrium holds for any positive value of ω.
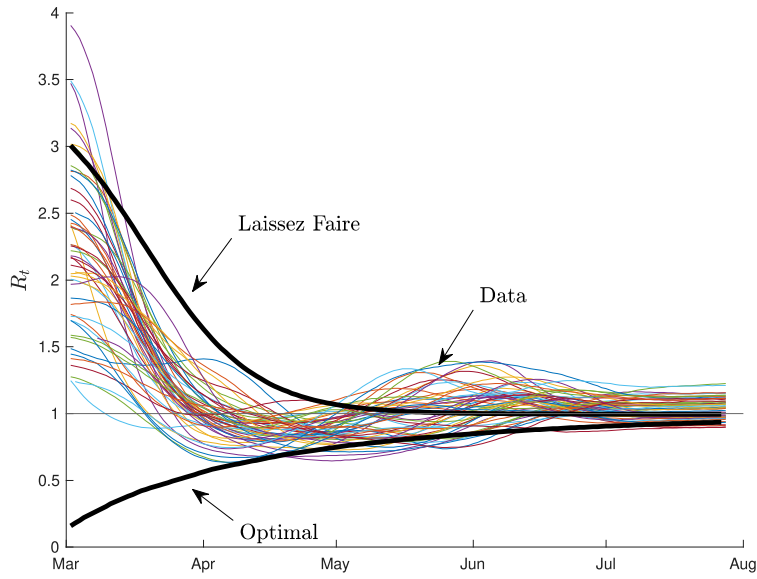
(a) Effective reproduction rate $R_{t,t}$. The thick lines depict the effective reproduction rate $(R_{t,t})$ according to the model under the optimal policy (OP) and laissez faire (LF) scenarios from beginning of March to end of July, 2020. The thin lines depict estimated $R_{t,t}$ for each US state.
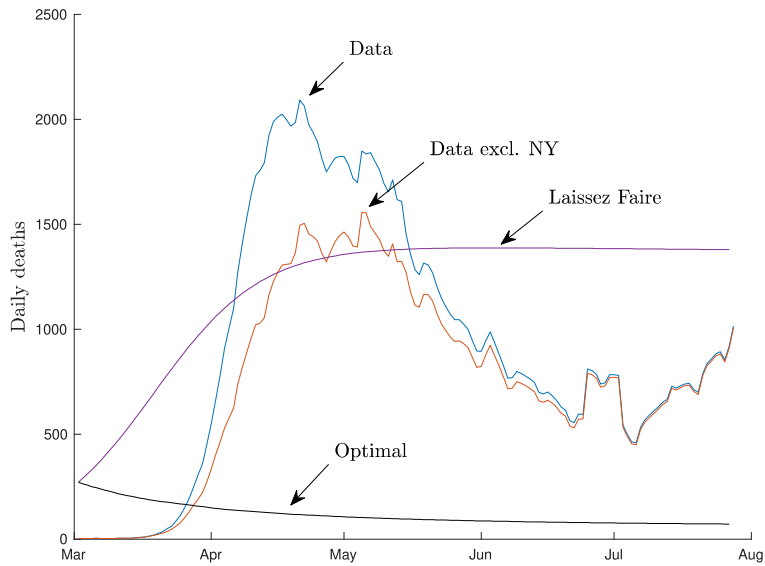
(b) Model-implied daily excess deaths under the optimal policy (OP) and laissez faire (LF) scenarios. The line "Data" corresponds to the daily excess deaths observed in the US over this period. The line "Data Excl NY" excludes New York from the computation of daily excess deaths.

Figure 2: Effective Reproduction rates and excess deaths when $\chi = 1$.
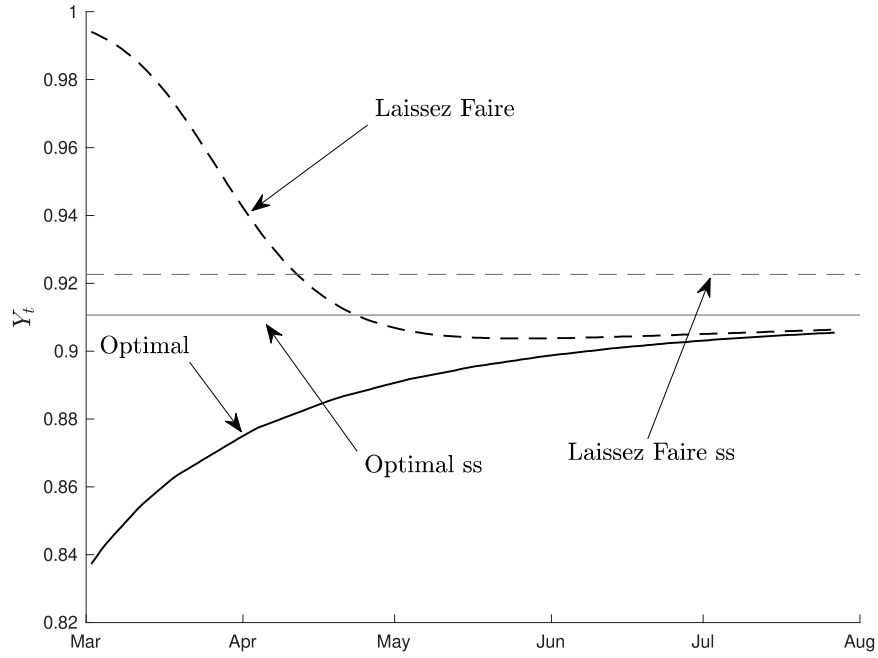
(a) Output Y_t under the optimal policy (OP) and laissez faire (LF) scenarios. The horizontal lines labelled "ss" reflect the steady-state values for the two scenarios.
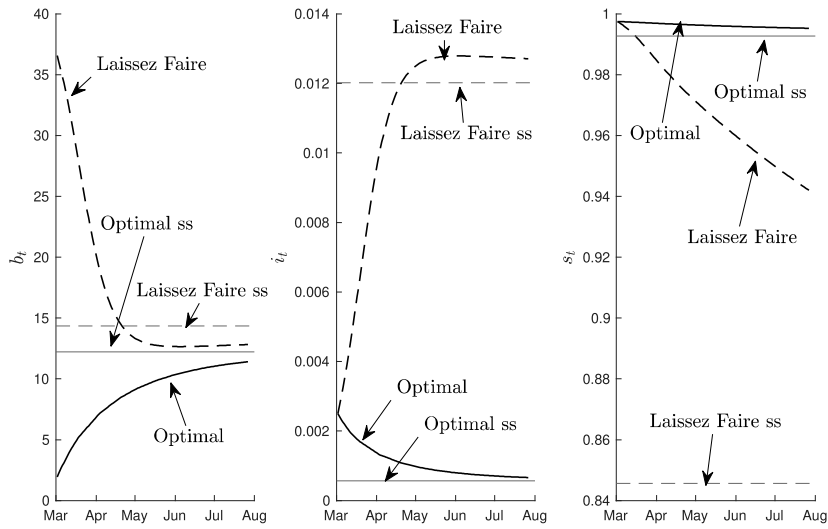
(b) b_t, i_t, s_t. The horizontal lines labelled "ss" reflect the steady-state values for the two scenarios.

Figure 3: Y_t, b_t, i_t, s_t when $\chi = 1$.

is $\frac{\delta}{365} \times Ni_0 = \frac{\delta}{365} \times (330 \times 10^6) \, i_0 = 270$, which together with $\delta = 0.12$ implies $i_0 = 0.0025$.

Figures 2 and 3 present our results for the baseline calibration ($\chi = 1$). In Figure 2(a), the optimal policy scenario (OP) shows the value of the effective reproduction rate, $\mathcal{R}_t$, when the contagion parameter β_t is the socially optimal value determined by the first-order condition in (34) at each point of time. The laissez-faire scenario (LF) shows the value of $\mathcal{R}_t$ when the contagion parameter β_t is the laissez-faire value determined by the first-order condition in (38) at each point of time. Figure 2(a) also shows empirical estimates of $\mathcal{R}_t$ for the 50 individual US states.[13] The most notable feature of Figure 2(a) is that scenario OP exhibits more aggressive policy to fight the disease than does scenario LF. Specifically, the values of $\mathcal{R}_t$, which reflect values of β_t, are lower in scenario OP than in scenario LF. In particular, during the first 4-6 weeks, $\mathcal{R}_t$ is substantially smaller than one in scenario OP and is substantially higher than one in scenario LF.

Figure 2(b) shows excess daily deaths. Under scenario OP, daily excess deaths initially decline and continue to decline throughout the 5 months shown. In contrast, under scenario LF, daily excess deaths spike upward abruptly, and after 2-3 months begin to decline very slowly. To illustrate the quantitative difference under the two scenarios, we find that on May 1, daily excess deaths are 107 under scenario OP and are 1,353 under scenario LF. For the first 6 weeks, the data for nationwide daily excess deaths resemble scenario LF. Thereafter, daily excess deaths decline rapidly, lying between those for LF and those for OP, as policies are implemented to combat the pandemic.

Figure 3(a) presents the evolution of output. Under scenario OP output drops immediately at the onset of the disease and rises towards its steady state value. The transition is quite fast; within two quarters, output is within one percent of its steady state value. Interestingly, the behavior of output under scenario LF is non-monotonic. Output declines by less than one percent in the first days of the pandemic. Then it declines to a level below its steady-state value, before rising quite slowly to its steady-state value. Figure 3(a) shows that around July the two scenarios imply essentially the same level of output, while Figure 2(b) shows that excess deaths are dramatically higher under LF than under OP.

---

[13]Source: http://rt.live. Data from http://epiforecasts.io imply quantitatively similar values.

Scenario OP shows the importance of sacrificing output early in the pandemic to prevent the stock of infections, i_t, from rising to an unmanageable level. After a few months, the stock of i_t can be maintained by choosing a stable β_t close to its steady state value. By contrast, under the LF scenario, the level of output remains higher than under OP, with the consequence that the infection stock grows to a far higher level under LF than under OP. After a few months, this higher level of i_t induces individuals to cut back their social and productive interactions, leading to a drop in output of about the same size as the drop in scenario OP at that point in time. While output drops by similar amounts after a few months in the two scenarios, the level of infection, i_t, and consequently daily excess deaths, are dramatically higher in scenario LF than in scenario OP. Therefore, the higher output in the early months of the pandemic under the LF scenario proves to be short-lived and comes at the cost of a high level of infection and excess deaths.

Table 1 presents the steady states for the OP and LF scenarios under different parametric assumptions. The baseline values of the parameters are the same as those used in Figures 2 and 3 ($\phi = 0.015$, $\mu = 0.01$, $\gamma = 12$, $\delta = 0.12$, $\omega = 50$, $\Delta = 0.09$, $\overline{\beta} = 42.473$, $\rho = 0.03$ and $\chi = 1.0$.). The three panels of the table use $\chi = 1.0$ (baseline), $\chi = 0.5$, and $\chi = 0.0$, respectively. For each of the 18 rows in the table, steady-state social distancing under LF is smaller than steady-state social distancing under OP, as reflected by the fact that β^* is higher, usually substantially higher, under LF than under OP. Therefore, steady-state infection, i^*, is higher under LF than under OP, since individual people ignore the infection externality they impose when infected. Also, since s^* is so much lower under LF than under OP, we can infer that r^* is higher under LF than OP. Thus, output per capita is higher under LF than under OP.

The three panels of the table illustrate the quantitative importance of including waning immunity, $\chi > 0$. The differences between the panel with $\chi = 1.0$ and the panel with $\chi = 0.5$, are swamped by the differences between the panel with $\chi = 0.5$ and the panel with $\chi = 0.0$. That is, even a little bit of waning immunity has an important quantitive impact. Interestingly, without waning immunity, output per capita in the LF scenario is essentially unchanged from its pre-pandemic level.

	Optimal						Laissez Faire				
	s^*	i^*	R_0^*	Y^*	β^*		s^*	i^*	R_0^*	Y^*	β^*
					$\chi = 1.0$						
$\omega = 50$ (Baseline)	99.27	0.06	1.01	0.91	12.22		84.57	1.20	1.18	0.92	14.35
$\omega = 20$	98.90	0.09	1.01	0.91	12.27		67.57	2.52	1.48	0.94	17.96
$\omega = 70$	99.40	0.05	1.01	0.91	12.21		88.69	0.88	1.13	0.92	13.68
$\Delta = 0.15$	98.70	0.10	1.01	0.85	12.30		73.37	2.07	1.36	0.90	16.54
$\Delta = 0.05$	99.61	0.03	1.00	0.95	12.18		91.74	0.64	1.09	0.95	13.23
$\rho = 0.05$	99.28	0.06	1.01	0.91	12.22		88.68	0.88	1.13	0.92	13.68
					$\chi = 0.5$						
$\omega = 50$	98.58	0.06	1.01	0.91	12.31		72.20	1.14	1.38	0.94	16.80
$\omega = 20$	97.84	0.09	1.02	0.91	12.40		51.32	1.99	1.95	0.96	23.64
$\omega = 70$	98.85	0.05	1.01	0.91	12.28		79.13	0.86	1.26	0.93	15.33
$\Delta = 0.15$	97.41	0.11	1.03	0.86	12.46		56.99	1.76	1.75	0.94	21.29
$\Delta = 0.05$	99.26	0.03	1.01	0.95	12.23		84.57	0.63	1.18	0.96	14.35
$\rho = 0.05$	98.59	0.06	1.01	0.91	12.31		79.13	0.86	1.26	0.93	15.33
					$\chi = 0.0$						
$\omega = 50$	62.81	0.05	1.59	0.97	19.32		30.22	0.09	3.31	1.00	40.16
$\omega = 20$	52.69	0.06	1.90	0.98	23.03		29.37	0.09	3.40	1.00	41.32
$\omega = 70$	67.95	0.04	1.47	0.96	17.86		30.80	0.09	3.25	1.00	39.40
$\Delta = 0.15$	49.03	0.06	2.04	0.97	24.75		29.54	0.09	3.38	1.00	41.07
$\Delta = 0.05$	77.64	0.03	1.29	0.97	15.63		31.61	0.08	3.16	1.00	38.39
$\rho = 0.05$	63.05	0.05	1.59	0.97	19.25		30.79	0.09	3.25	1.00	39.41

Table 1: Steady state values of $s^*, i^*, R_0^*, Y^*, \beta^*$ for the optimal and "laissez faire" equilibria under alternative parametric assumptions. s^* and i^* are expressed in percentage points.

6 Optimal Social Distancing and Vaccination Policies

In Section 3, we examined optimal social distancing in the absence of a vaccine. In this section, we extend the analysis of optimal policy to include optimal vaccination rates as well as optimal social distancing. We assume that

$$\lambda \leq \lambda_{\max}, \tag{39}$$

where $\left(\frac{\bar{\beta}}{\beta_c} - 1\right)(\phi + \chi) < \lambda_{\max} < \infty$. Importantly, the lower bound on $\lambda_{\max}$ allows λ to be large enough to achieve a DFE steady state, even when there is no social distancing so

29

$\beta = \overline{\beta}$.[14]

We derive optimal policies in a dynamic framework and then focus on the steady state of an economy following optimal policies. The planner's maximization problem is

$$\max_{\beta_u, \lambda_u \leq \lambda_{\max} \text{ for } u \geq t} E_t \left\{ \int_t^{\infty} N_u e^{-\rho u} \left[Y\left(\beta_u, s_u, r_u\right) - \omega \delta i_u - \psi \lambda_u \left(s_u + \zeta r_u\right) \right] du \right\}, \qquad (40)$$

where $\psi > 0$ and $\psi \lambda_u \left(s_u + \zeta r_u\right) N_u$ is the real cost of vaccinating the vaccine-eligible population $\left(s_u + \zeta r_u\right) N_u$ at rate λ_u at time $u \geq t$.

The Bellman equation associated with the maximization problem in equation (40) is

$$\left(\rho - g\left(i\right)\right) V\left(s, i, r\right) = \max_{\beta, \lambda \leq \lambda_{\max}} \left\{ \begin{array}{c} Y\left(\beta, s, r\right) - \omega \delta i - \psi \lambda \left(s + \zeta r\right) \\ + V_s \frac{ds}{dt} + V_i \frac{di}{dt} + V_r \frac{dr}{dt} \end{array} \right\}. \qquad (41)$$

Appendix A derives expressions for the dynamic behavior of $V_j(s, i, r)$, $j \in \{s, i, r\}$, by differentiating both sides of equation (41) with respect to j and using $\frac{dV_j}{dt} = V_{js} \frac{ds}{dt} + V_{ji} \frac{di}{dt} + V_{jr} \frac{dr}{dt}$. As shown in that appendix,

$$\left(\rho + \mu\right) V_s = y\left(\beta\right) - \psi \lambda + \frac{dV_s}{dt} + \beta i \left(V_i - V_s\right) + \lambda \left(V_r - V_s\right), \qquad (42)$$

which is the same as equation (31), except for three straightforward changes. First, the hazard rate of the future arrival of the vaccine, p, that appears in equation (31) is absent from equation (42) because the vaccine has already arrived. Second, the right hand side of equation (42) subtracts the cost of vaccinating a susceptible person, $\psi \lambda$, which is not present in equation (31). Third, the right hand side of equation (42) includes the hazard, λ, that a susceptible person is successfully vaccinated and joins the pool of recovered people, multiplied by the value, $V_r - V_s$, of that transition.

Appendix A also derives the dynamic equations for V_i and V_r. These equations are identical to equations (32) and (33), respectively, once one sets $p = 0$ in equations (32) and

---

[14]Proposition 5 implies, that if $\beta = \overline{\beta}$ and $\lambda = \left(\frac{\overline{\beta}}{\beta_c} - 1\right)\left(\phi + \chi\right)$, then $i^* = 0$.

(33). For convenience, we repeat these equations here

$$\left(\rho + \mu + \delta\right) V_i = -\omega\delta + \frac{dV_i}{dt} + \gamma\left(V_r - V_i\right) - \delta\left[V - \left(sV_s + iV_i + rV_r\right)\right] - \beta s\left(V_s - V_i\right) \quad (43)$$

and

$$\left(\rho + \mu\right) V_r = y\left(\beta_r\right) - \zeta\psi\lambda + \frac{dV_r}{dt} + \chi\left(V_s - V_r\right). \quad (44)$$

The first-order conditions for optimal β and optimal λ are obtained by differentiating the right hand side of equation (41) with respect to β and λ, respectively. The first-order condition for β is

$$\left(s + \zeta r\right) y'\left(\beta\right) = \left(V_s - V_i\right) si, \quad (45)$$

which is the same equation (34), and the first-order condition for optimal λ is

$$\left(s + \zeta r\right) \psi \leq \left(V_r - V_s\right) s. \quad (46)$$

The inequality in equation (46) arises from the restriction $\lambda \leq \lambda_{\max}$.

6.1 Optimal Social Distancing and Optimal Vaccination in the Steady State

In the steady state, $\frac{ds}{dt} = \frac{di}{dt} = \frac{dr}{dt} = \frac{dV_s}{dt} = \frac{dV_i}{dt} = \frac{dV_r}{dt} = 0$. In this subsection we analyze the optimal combination of β and λ in the steady state, which we denote as $(\beta_{OP}^*, \lambda_{OP}^*)$. We begin the analysis of optimal policy in the steady state with two lemmas that, taken together, imply that $(\beta_{OP}^*, \lambda_{OP}^*)$ must lie in the triangle ABC in Figure 1. All (β, λ) pairs on the line segment $\overline{AB}$, which is the northwest boundary of this triangular region, are associated with DFE steady states. All (β, λ) pairs in the interior of ABC, as well as the boundaries $\overline{AC}$ (excluding Point A) and $\overline{BC}$ (excluding Point B), are associated with EE steady states. We then show that $(\beta_{OP}^*, \lambda_{OP}^*)$ cannot be on $\overline{AB}$, thereby proving Proposition 10 below, which

31

states that the optimal policy leads to a steady state that is EE. Since the proof of Lemma 2 is instructive, we include it in the text.

Lemma 2 $(\beta_{OP}^*, \lambda_{OP}^*)$ *cannot lie to the left of the piecewise-linear boundary ABD in Figure 1.*

Proof of Lemma 2. Suppose, contrary to what is to be proved, that $(\beta_{OP}^*, \lambda_{OP}^*)$ is located at an arbitrary point, say G, to the left of ABD. Now consider Point H in Figure 1, which is the point on ABD that is directly to the right of Point G, with a higher value of β and the same value of λ as at Point G. The instantaneous flow of aggregate output, $Y(\beta^*, s^*, r^*)$, is higher at Point H than at Point G, because $Y_\beta > 0$ and Point H has a higher value of β^* and the same values of s^* and r^* as Point G. Thus, Point G cannot represent the (β, λ) pair with the highest steady-state value of the objective function. Moreover, this finding goes beyond a simple comparative steady-state analysis that ignores the transition path from the steady state at Point G to the steady state at Point H. The comparison described here fully takes account of the transition path because *there is no transition path from the steady state at Point G to the steady state at Point H.* The values of $s^* = 1 - \frac{\lambda}{\phi+\chi+\lambda}$, $i^* = 0$, and $r^* = \frac{\lambda}{\phi+\chi+\lambda}$ are the same at Points G and H, so the new steady state is reached immediately. Therefore, no point to the left of ABD can be optimal. ∎

Lemma 3 $(\beta_{OP}^*, \lambda_{OP}^*)$ *cannot lie on the half-open line segment $\overline{BD}$ (which includes point D but excludes point B) in Figure 1.*

Lemmas 2 and 3 imply that $(\beta_{OP}^*, \lambda_{OP}^*)$ lies in the triangle $\overline{ABC}$, either on the boundary or in the interior. Points along the boundary $\overline{AB}$ are associated with DFE steady states; all other points in ABC, including the interior of the triangle as well as the half-open line segment $\overline{AC}$ (excluding point A) and the half-open line segment $\overline{BC}$ (excluding Point B) are associated with EE steady states.

The following equation, which is derived by subtracting equation (42) from equation (44) and setting $\frac{dV_s}{dt} = \frac{dV_r}{dt} = 0$, is useful in proving Proposition 10 below

$$(\rho + \mu + \chi + \lambda^*)(V_r^* - V_s^*) = y(\beta_r^*) - y(\beta^*) + (1 - \zeta)\psi\lambda^* + \beta i^*(V_s^* - V_i^*). \qquad (47)$$

32

Proposition 10 *Assume that a vaccine is available and let $\lambda_h(\beta)$ be defined as in Proposition 5. Assume that $\frac{y(\overline{\beta})}{\psi} > \max[\lambda_h(\overline{\beta}), \gamma + \delta + \rho + \mu + \chi]$. Under optimal social distancing and vaccination policies, the steady state is EE. The optimal steady-state values of β and λ, β_{OP}^* and λ_{OP}^*, respectively, satisfy $\beta_c < \beta_{OP}^* < \overline{\beta}$ and $0 \leq \lambda_{OP}^* < \lambda_h(\beta_{OP}^*)$.*

Proposition 10 states that even when a vaccine is available, the optimal steady state is EE. Since $\beta_{OP}^* < \overline{\beta}$, there will be some degree of social distancing in the optimal steady state, but since $\beta_{OP}^* > \beta_c$, social distancing alone will not be intense enough to eliminate the disease in the steady state. While it is feasible for the vaccine to eliminate the disease for any degree of social distancing by setting $\lambda^* = \lambda_h(\beta_{OP}^*)$, Proposition 10 states that $\lambda_{OP}^* < \lambda_h(\beta_{OP}^*)$.

We have shown that $(\beta_{OP}^*, \lambda_{OP}^*)$ must lie in the triangle ABC in Figure 1, which implies that $\lambda_{OP}^* \leq \lambda_h(\overline{\beta}) < \lambda_{\max}$. Therefore, in the steady state, the constraint $\lambda \leq \lambda_{\max}$ in equation (39) is not binding[15] so the first-order condition for ψ in equation (46) holds as an equality $\psi(s + \zeta r) = (V_r - V_s) s$. We can substitute this first-order condition for λ^* and the first-order condition for β^* from equation (45) into equation (47) to obtain

$$(\rho + \mu + \chi + \lambda^*)\left(1 + \zeta\frac{r^*}{s^*}\right)\psi = y(\beta_r^*) - y(\beta_{OP}^*) + (1 - \zeta)\psi\lambda^* + \left(1 + \zeta\frac{r^*}{s^*}\right)\beta_{OP}^* y'(\beta_{OP}^*) \quad (48)$$

Equation (48) relates the optimal steady-state values of λ and β without using the value function, V, or its derivatives, V_s, V_i, and V_r. Nevertheless, equation (48) depends on the endogenous variables s^* and r^*. However, we can use the definition $\beta_r \equiv \zeta\beta + (1 - \zeta)\overline{\beta}$ from equation (22) to obtain a relationship between λ^* and β^* that involves only fundamental parameters and the specification of the production function $y(\beta)$. Specifically,[16]

$$(\rho + \mu + \chi + \zeta\lambda^*)\psi = (1 - \zeta)\left[y(\overline{\beta}) - y(\beta_{OP}^*)\right] + \beta_{OP}^* y'(\beta_{OP}^*). \quad (49)$$

---

[15]This statement leaves open the possibility that early in the pandemic, when almost everyone is susceptible to the disease, it may be optimal to vaccinate as quickly as possible. That is, the constraint $\lambda \leq \lambda_{\max}$ might bind outside of the steady state.

[16]If $\zeta = 0$, then $\beta_r = \overline{\beta}$ and equation (48) becomes $(\rho + \mu + \chi)\psi = y(\overline{\beta}) - y(\beta_{OP}^*) + \beta_{OP}^* y'(\beta_{OP}^*)$. If $\zeta = 1$, then $\beta_r^* = \beta_{OP}^*$ and equation (48) becomes $(\rho + \mu + \chi + \lambda^*)\psi = \beta_{OP}^* y'(\beta_{OP}^*)$.

	i^*	R_0^*	Y^*	β^*	λ^*
			$\chi = 1.0$		
$\omega = 50$ (Baseline)	0.000184	3.394	0.999950	41.186	2.430
$\omega = 20$	0.000266	3.394	0.999950	41.186	2.430
$\omega = 70$	0.000152	3.394	0.999951	41.186	2.430
$\Delta = 0.15$	0.000186	3.436	0.999970	41.699	2.473
$\Delta = 0.05$	0.000178	3.310	0.999911	40.166	2.344
$\rho = 0.05$	0.000406	3.393	0.999948	41.179	2.429
$\psi = 0.002$	0.000123	3.429	0.999978	41.614	2.466
$\psi = 0.001$	0.000062	3.465	0.999994	42.042	2.501
			$\chi = 0.5$		
$\omega = 50$	0.000095	3.446	0.999987	41.814	1.260
$\omega = 20$	0.000138	3.446	0.999986	41.814	1.259
$\omega = 70$	0.000079	3.446	0.999987	41.814	1.260
$\Delta = 0.15$	0.000096	3.467	0.999992	42.077	1.271
$\Delta = 0.05$	0.000094	3.402	0.999977	41.289	1.237
$\rho = 0.05$	0.000212	3.445	0.999985	41.807	1.259
$\psi = 0.002$	0.000064	3.464	0.999994	42.033	1.269
$\psi = 0.001$	0.000032	3.482	0.999998	42.253	1.278
			$\chi = 0.0$		
$\omega = 50$	0.000004	3.498	1.000000	42.444	0.037
$\omega = 20$	0.000006	3.498	1.000000	42.445	0.037
$\omega = 70$	0.000003	3.498	1.000000	42.445	0.037
$\Delta = 0.15$	0.000004	3.499	1.000000	42.456	0.037
$\Delta = 0.05$	0.000004	3.496	1.000000	42.422	0.037
$\rho = 0.05$	0.000011	3.497	1.000000	42.437	0.037
$\psi = 0.002$	0.000003	3.498	1.000000	42.454	0.037
$\psi = 0.001$	0.000001	3.499	1.000000	42.463	0.037

Table 2: Steady-state values of $i^*, R_0^*, Y^*, \beta^*, \lambda^*$ in the presence of vaccination and with $\zeta = 1$. i^* is expressed in percentage points.

Equation (49) provides a remarkably simple illustration of our finding that the steady state under the optimal combination of social distancing and vaccination policy is endemic. The left hand side of equation (49) is strictly positive so the right hand side must also be strictly positive. Since $y(\beta) \equiv 1$ for $\beta \geq \overline{\beta}$, the fact that the right hand side of equation (49) is positive implies $\beta_{OP}^* < \overline{\beta}$. The only reason for social distancing, represented by $\beta_{OP}^* < \overline{\beta}$, is because $i^* > 0$ (i.e., the disease remains present in the population in the steady state).

Since λ^* is finite, $\lim_{\psi \to 0} (\rho + \mu + \chi + \lambda^*) \psi = 0$. Therefore, equation (49) implies

that $\lim_{\psi \to 0} \left[(1 - \zeta) \left[y \left(\overline{\beta} \right) - y \left(\beta_{OP}^* \right) \right] + \beta_{OP}^* y' \left(\beta_{OP}^* \right) \right] = 0$, which implies $\lim_{\psi \to 0} \beta_{OP}^* = \overline{\beta}$.[17] That is, when the cost of the vaccine is arbitrarily small, the optimal amount of social distancing is arbitrarily small because the vaccine essentially bears the total burden of reducing infections.

Equation (49) is linear in λ^* and since $y(\beta)$ is quadratic in β, equation (49) is quadratic in β^*. Specifically, for $\beta \leq \overline{\beta}$, the production function is $y(\beta) = 1 - \alpha \left(\overline{\beta} - \beta \right)^2$ so $\beta y'(\beta) = 2\alpha\beta \left(\overline{\beta} - \beta \right)$. Therefore, equation (49) can be rewritten as

$$(\rho + \mu + \chi + \zeta\lambda^*)\psi = (1 - \zeta)\alpha \left(\overline{\beta} - \beta_{OP}^* \right)^2 + 2\alpha\beta_{OP}^* \left(\overline{\beta} - \beta_{OP}^* \right), \quad \text{for } \beta^* \leq \overline{\beta}. \quad (50)$$

When $\zeta = 0$, equation (50) is a quadratic equation in β_{OP}^* and is independent of λ^*. We solve this equation to obtain[18]

$$1 - \left(\frac{\beta_{OP}^*}{\overline{\beta}} \right)^2 = \frac{\rho + \mu + \chi}{\alpha\overline{\beta}^2} \frac{\psi}{y\left(\overline{\beta} \right)}. \quad (51)$$

The left hand side of equation (51) is a (nonlinear) measure of social distancing. In our baseline calculations, $\rho = 0.03$ per year, $\mu = 0.01$ per year, $\chi = 1$ per year, $\overline{\beta} = 42.473$, $\alpha = 9.779 \times 10^{-5}$ and hence $\alpha\overline{\beta}^2 = 0.1764$, so the first of the two factors on the right hand side of equation (51) is $\frac{\rho + \mu + \chi}{\alpha\overline{\beta}^2} = \frac{1.04}{0.1764} \approx 5.90$. The second factor on the right hand side of equation (51), $\frac{\psi}{y\left(\overline{\beta} \right)}$, is the marginal cost of the vaccine divided by output per capita. In our baseline calibration we use $\frac{\psi}{y\left(\overline{\beta} \right)} = 0.003$.[19] Therefore, the right hand side of equation (51) is $5.90 \times 0.003 \approx 0.02$. Hence, equation (51) implies that $\left(\frac{\beta_{OP}^*}{\overline{\beta}} \right)^2 \approx 0.98$, which implies $\frac{\beta_{OP}^*}{\overline{\beta}} \approx 0.99$, and hence $\left(1 - \frac{\beta_{OP}^*}{\overline{\beta}} \right)^2 \approx (0.01)^2 = 10^{-4}$. Therefore, the loss in output per

---

[17]When $\zeta = 1$, $\beta_{OP}^* = 0$ also satisfies $\left[(1 - \zeta) \left[y \left(\overline{\beta} \right) - y \left(\beta_{OP}^* \right) \right] + \beta_{OP}^* y' \left(\beta_{OP}^* \right) \right] = 0$, but violates $\beta_{OP}^* \geq \beta_c > 0$.

[18]With $\zeta = 0$, equation (50) becomes $(\rho + \mu + \chi)\psi = \alpha \left(\overline{\beta} - \beta_{OP}^* \right)^2 + 2\alpha\beta_{OP}^* \left(\overline{\beta} - \beta_{OP}^* \right) = \alpha \left(\overline{\beta} - \beta_{OP}^* \right) \left(\overline{\beta} + \beta_{OP}^* \right) = \alpha \left(\overline{\beta}^2 - \beta_{OP}^{*2} \right)$. Dividing this equation by $\overline{\beta}^2$, and recalling the normalization $y \left(\overline{\beta} \right) = 1$ yields $\frac{\rho + \mu + \chi}{\overline{\beta}^2} \frac{\psi}{\alpha y \left(\overline{\beta} \right)} = 1 - \left(\frac{\beta}{\overline{\beta}} \right)^2$.

[19]For an average income of $60,000 per year, this implies $\psi = 180. While the manufacturing cost of a vaccine is only a small fraction of that number, vaccination also involves other costs (travel time to and from vaccination centers, administrative costs, healthcare-professional time, venue rentals etc.)

capita, $y\left(\overline{\beta}\right) - y\left(\beta_{OP}^*\right) = \alpha \overline{\beta}^2 \left(1 - \frac{\beta_{OP}^*}{\overline{\beta}}\right)^2$, is approximately 0.1764×10^{-4}, which is less than 20% of one basis point.

In the case in which $\zeta = 1$, equation (50) depends on λ^* as well as on β_{OP}^*. In this case, it is not possible to solve for β^* analytically. Table 2 provides numerical results for the case with $\zeta = 1$ for various parameter assumptions. As in the case with $\zeta = 0$, Y^* is within one basis point of $y(\overline{\beta}) = 1$. For a fixed value of χ, there are no notable differences across the various parameter constellations. The most interesting variation is across different values of χ. In particular, the optimal vaccination rate is approximately equal to $2.5(\phi + \chi)$, which accounts for the differences in λ^* as χ changes from 1.0 to 0.5 to 0. [20]

7 Concluding Remarks

In the epidemiology literature, there is a sharp contrast between disease-free equilibria and endemic equilibria. The major theoretical result of our paper is that optimal public health policy leads to an endemic steady state even if an effective vaccine is available. This result is consistent with the quoted statements of Drs. Heyman and Fauci at the beginning of the paper. However, the planner's problem that underlies our finding provides a framework to calculate the loss in steady-state welfare caused by the ongoing nature of the disease and the policies to manage it.

Equation (40) implies that under optimal policy, the flow of welfare per capita in the steady state is $Y\left(\beta_{OP}^*, s^*, r^*\right) - \omega \delta i^* - \psi \lambda^* \left(s^* + \zeta r^*\right)$. If $\zeta = 1$, which is the case in which vaccines are given to both susceptible people and recovered people, the steady-state flow of welfare is

$$Y^* - \omega \delta i^* - \psi \lambda^* \left(1 - i^*\right), \tag{52}$$

---

[20]The following approximation offers a transparent view of the impacts of ψ and χ on the determination of λ^*. With optimal use of the vaccine, $\beta^* \approx \overline{\beta}$ and $i^* \approx 0$ so $\lambda^* \approx \lambda_h\left(\overline{\beta}\right) = \left(\frac{\overline{\beta}}{\beta_c} - 1\right)(\phi + \chi)$. In Section 5, we calculated $\overline{\beta}$ as $(\beta_c - \delta i)\overline{R}_0 \approx \beta_c \overline{R}_0$, where $\overline{R}_0$ is the value of R_0 observed on the initial day of the pandemic when β is still at its pre-pandemic level, $\overline{\beta}$, and i is approximately zero. Substituting $\beta_c \overline{R}_0$ for $\overline{\beta}$ in the expression for $\lambda_h(\overline{\beta})$ yields $\lambda_h(\overline{\beta}) = (\overline{R}_0 - 1)(\phi + \chi)$. Using $\overline{R}_0 = 3.5$ from Section 5, $\lambda_h(\overline{\beta}) = 2.5(\phi + \chi)$.

| | Pre-vaccine | | |
χ	$1\text{-}Y^*$	$\omega\delta i^*$	L^*
1	900	36	936
0.5	900	36	936
0	300	30	330

| | Cum-vaccine | | | |
χ	$1\text{-}Y^*$	$\omega\delta i^*$	$\psi(1-i^*)\lambda^*$	L^*
1	0.00	0.12	72.89	73.01
0.5	0.00	0.06	37.78	37.84
0	0.00	0.00	1.13	1.13

Table 3: Loss in Steady-State Flow of Welfare. This table shows the components of the loss in steady-state welfare, and the sum of these components, denoted L^*. All numerical entries are expressed in basis points.

where Y^* is $Y\left(\beta_{OP}^*, s^*, r^*\right)$. As a benchmark for comparison, if the disease had never appeared in the first place, the steady-state flow of welfare would be $y\left(\overline{\beta}\right) = 1$ because β^* would be $\overline{\beta}$ and i^* and λ^* would be zero. Therefore, the reduction in the steady-state flow of welfare resulting from the disease and optimal public health policy is 1 minus the expression in equation (52). Letting L^* be this loss,

$$L^* = (1 - Y^*) + \omega\delta i^* + \psi\lambda^* (1 - i^*). \tag{53}$$

Equation (53) illustrates the three components of the loss, namely, the loss in output per capita, the loss in welfare arising from disease-induced deaths, and the cost of administering vaccinations. Table 3 displays, for the pre-vaccine model and the cum-vaccine model, the components of L^* for the baseline calculations when $\chi \in \{0, 0.5, 1.0\}$ in Tables 1 and 2, respectively. To avoid long strings of zeroes after the decimal point in some entries of the table, all of the entries are expressed in basis points. Several conclusions are evident in Table 3. First, without a vaccine, the loss in welfare, L^*, is substantial; when $\chi = 1$, the loss in welfare is 9.36% of the never-disease benchmark. Second, in the absence of a vaccine, with social distancing being the only measure to reduce the spread of the disease, the loss in output per capita contributes far more to L^* than do disease-induced deaths; when $\chi = 1$, optimal social distancing is intense enough to reduce output per capita by 9.00%

while disease-induced deaths reduce L^* by 0.36%. Third, the introduction of vaccines to the toolkit of public health policy reduces L^* by more than 90% for $\chi = 1$ and by more than 99% for $\chi = 0$. Fourth, consistent with the analytic findings discussed above for the case with $\zeta = 0$, the introduction of vaccines virtually eliminates the need for social distancing in the steady state, as evidenced by the finding that the reduction in steady state output per capita is smaller than one percent of one basis point. Fifth, in the presence of vaccines, more than 99% of L^* is due to the cost of the vaccinating the population; with $\chi = 1$, the cost of optimal vaccination reduces welfare by 0.73%. With aggregate output per capita virtually unaffected by the disease in the steady state, the cost of vaccination amounts to a redirection of 0.73% of output away from other goods and services.

Our overall conclusion is that optimal social distancing and vaccination lead to an endemic steady state, regardless of the welfare cost assigned to the loss of a life from the disease. To give our model the greatest opportunity to find that optimal policy might lead to a disease-free steady state, we incorporated total compliance by the public with mandated social distancing and vaccination guidelines. Nevertheless, the propositions we derived state that optimal policy leads to an endemic state, with or without a vaccine. For the range of parameter values examined in the quantitative analysis of our model, the steady-state flow of welfare under total compliance with optimal public health guidelines can be much as 0.73% below the steady-state flow of welfare in the never-disease steady state. Taking account of noncompliance with social distancing mandates and the stated refusal by many people to get vaccines could substantially increase the degree of infection and deaths, reinforcing the major result of this paper.

References

Acemoglu, D., V. Chernozhukov, I. Werning, and M. D. Whinston (2020). Optimal Targeted Lockdowns in a Multi-Group SIR Model. Working Paper, MIT.

Alvarez, F., D. Argente, and F. Lippi (2020). A Simple Planning Problem for COVID-19 Lockdown. Becker Friedman Institute Working paper 2020-34.

Atkeson, A. G. (2020). What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios. National Bureau of Economic Research, Working Paper, 26867.

Avery, C., W. Bossert, A. Clark, G. Ellison, and S. F. Ellison (2020). An economist's guide to epidemiology models of infectious disease. *The Journal of Economic Perspectives 34*(4), 79–104.

Bailey, N. (1975). *The Mathematical Theory of Infectious Diseases and its Applications.* London: Griffin.

Barrett, S. and M. Hoel (2007). Optimal disease eradication. *Environment and Development Economics 12*(5), 627–652.

Berger, D. W., K. F. Herkenhoff, and S. Mongey (2020). An SEIR infectious disease model with testing and conditional quarantine. National Bureau of Economic Research, Working Paper, 26901.

Bethune, Z. A. and A. Korinek (2020). Covid-19 infection externalities: Trading off lives vs. livelihoods. Working Paper 27009, National Bureau of Economic Research.

Cenesiz, M. A. and L. Guimarães (2020). Covid-19: What if immunity wanes?

Cochrane, J. C. (2020). An SIR model with behavior. "The Grumpy Economist" Blog, Monday May 4, 2020.

Eichenbaum, M. S., S. Rebelo, and M. Trabandt (2020). The Macroeconomics of Epidemics. National Bureau of Economic Research, Working Paper 26882.

Farboodi, M., G. Jarosch, and R. Shimer (2020). Internal and External Effects of Social Distancing in a Pandemic. NBER Working Papers 27059, National Bureau of Economic Research, Inc.

Fernandez-Villaverde, J. and C. I. Jones (2020). Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities. Working Paper, Stanford GSB and University of Pennsylvania.

Garibaldi, P., E. R. Moen, and C. Pissarides (2020). Static and dynamic inefficiencies in an optimizing model of epidemics. CEPR Working Papers 15439, Centre for Economic Policy Research.

Geoffard, P.-Y. and T. Philipson (1997). Disease eradication: Private versus public vaccination. *The American Economic Review 87*(1), 222–230.

Gersovitz, M. and J. S. Hammer (2003). The Economical Control of Infectious Diseases. *The Economic Journal 114*(492), 1–27.

Giannitsarou, C., S. Kissler, and F. Toxvaerd (2020). Waning Immunity and the Second Wave: Some Projections for SARS-CoV-2. CEPR Discussion Papers 14852, C.E.P.R. Discussion Papers.

Goldman, S. M. and J. Lightwood (2002). Cost Optimization in the SIS Model of Infectious Disease with Treatment. *The B.E. Journal of Economic Analysis & Policy 2*(1), 1–24.

Gonzalez-Eiras, M. and D. Niepelt (2020a). On the optimal 'lockdown' during an epidemic. *Covid Economics* (7), 68–87.

Gonzalez-Eiras, M. and D. Niepelt (2020b). Optimally controlling an epidemic. CEPR Discussion Papers 15541, C.E.P.R. Discussion Papers.

Hall, R. E., C. I. Jones, and P. J. Klenow (2020). Trading Off Consumption and COVID-19 Deaths. *Federal Reserve Bank of Minneapolis Quarterly Review 42*(1).

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review 42*(4), 599–653.

Jones, C. J., T. Philippon, and V. Venkateswaran (2020). Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home. NBER Working Papers 26984, National Bureau of Economic Research, Inc.

Kermack, W. O., A. G. McKendrick, and G. T. Walker (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 115*(772), 700–721.

Piguillem, F. and L. Shi (2020). Optimal Covid-19 Quarantine and Testing Policies. CEPR Discussion Paper No. DP14613.

Toxvaerd, F. (2020). Equilibrium Social Distancing. Cambridge-INET Working Paper Series No: 2020/08.

Yaari, M. E. (1965). Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *Review of Economic Studies 32*(2), 137–150.

Appendix

A Optimal Social Distancing and Vaccination Policies

In this appendix, we derive optimal policies both when social distancing is the only policy tool and when both vaccination and social distancing are policy tools. We can efficiently derive these optimal policies in the context of a single problem that encompasses both the pre-vaccine analysis in Section 3 and the cum-vaccine analysis in Section 6.

As a first step, use $\mu + g = \phi - \delta i$ to write equations (6) - (8) as

$$\frac{ds}{dt} = \phi + \chi r - \beta s i - (\phi - \delta i + \lambda) s \tag{A.1}$$

$$\frac{di}{dt} = [\beta s - (\delta + \gamma + \phi - \delta i)] i \tag{A.2}$$

$$\frac{dr}{dt} = \gamma i + \lambda s - (\phi - \delta i + \chi) r. \tag{A.3}$$

Consider the following Bellman equation that nests the Bellman equations (30) and (41) for the pre-vaccine and cum-vaccine models, respectively

$$(\rho + p - g(i)) V(s, i, r) = \max_{\beta, \lambda \leq \lambda_{\max}} \left\{ \begin{array}{c} s y(\beta) + r y(\beta_r) - \omega \delta i - \psi \lambda (s + \zeta r) \\ + V_s \frac{ds}{dt} + V_i \frac{di}{dt} + V_r \frac{dr}{dt} + p \widetilde{V} \end{array} \right\}. \tag{A.4}$$

Note when $\lambda \equiv 0$, the Bellman equation (A.4) is identical to Bellman equation (30) and when $p = 0$ the Bellman equation (A.4) is identical to Bellman equation (41).

Differentiate both sides of (A.4) with respect to s and use $\frac{dV_s}{dt} = V_{ss} \frac{ds}{dt} + V_{si} \frac{di}{dt} + V_{sr} \frac{dr}{dt}$ and

(A.1) - (A.3) to obtain

$$(\rho + p - g(i)) V_s = y(\beta) - \psi\lambda + \frac{dV_s}{dt} - (\phi + (\beta - \delta) i + \lambda) V_s + \beta i V_i + \lambda V_r. \qquad \text{(A.5)}$$

Use $g(i) = \phi - \mu - \delta i$ and define

$$\sigma \equiv \rho + \mu + p \qquad \text{(A.6)}$$

to obtain

$$\sigma V_s = y(\beta) - \psi\lambda + \frac{dV_s}{dt} + \beta i (V_i - V_s) + \lambda (V_r - V_s). \qquad \text{(A.7)}$$

Differentiate both sides of (A.4) with respect to i and use $\frac{dV_i}{dt} = V_{is}\frac{ds}{dt} + V_{ii}\frac{di}{dt} + V_{ir}\frac{dr}{dt}$, $g(i) = \phi - \mu - \delta i$ and (A.1) - (A.3) to obtain

$$\sigma V_i + \delta V = -\omega\delta + \frac{dV_i}{dt} - (\beta - \delta) s V_s + [\beta s - (\delta + \gamma - \delta i)] V_i + (\gamma + \delta r) V_r, \qquad \text{(A.8)}$$

which can be rearranged to obtain

$$(\sigma + \delta) V_i = -\omega\delta + \frac{dV_i}{dt} + \beta s (V_i - V_s) + \gamma (V_r - V_i) - \delta [V - (s V_s + i V_i + r V_r)]. \qquad \text{(A.9)}$$

Differentiate both sides of (A.4) with respect to r and use $\frac{dV_r}{dt} = V_{rs}\frac{ds}{dt} + V_{ri}\frac{di}{dt} + V_{rr}\frac{dr}{dt}$, $g(i) = \phi - \mu - \delta i$, and equations (A.1) - (A.3) to obtain

$$\sigma V_r = y(\beta_r) - \zeta\psi\lambda + \frac{dV_r}{dt} + \chi (V_s - V_r). \qquad \text{(A.10)}$$

Equations (A.7), (A.9) and (A.10) can be written as a first-order system of nonhomoge-

neous linear ordinary differential equations with nonconstant coefficients

$$
\begin{bmatrix}
\frac{dV_s}{dt} \\[2mm]
\frac{dV_i}{dt} \\[2mm]
\frac{dV_r}{dt}
\end{bmatrix}
= A
\begin{bmatrix}
V_s \\[2mm]
V_i \\[2mm]
V_r
\end{bmatrix}
- b
\tag{A.11}
$$

where

$$
A \equiv
\begin{bmatrix}
\sigma + \beta i + \lambda & -\beta i & -\lambda \\[2mm]
(\beta - \delta)\,s & \sigma + \delta + \gamma - \beta s - \delta i & -(\gamma + \delta r) \\[2mm]
-\chi & 0 & \sigma + \chi
\end{bmatrix},
\tag{A.12}
$$

and

$$
b \equiv
\begin{bmatrix}
y(\beta) - \psi\lambda \\[2mm]
-\delta(V + \omega) \\[2mm]
y(\beta_r) - \zeta\psi\lambda
\end{bmatrix}.
\tag{A.13}
$$

A.1 Steady State

Now confine attention to the steady state so that $\frac{ds}{dt} = \frac{di}{dt} = \frac{dr}{dt} = \frac{dV_s}{dt} = \frac{dV_i}{dt} = \frac{dV_r}{dt} = 0$. Set $\frac{ds}{dt} = \frac{di}{dt} = \frac{dr}{dt} = 0$ in equation (A.4) and use the fact, which we verify below, that V_s^*, V_i^*, and V_r^* are finite to obtain $(\rho + p - g(i^*))\,V^* = s^*\,[y(\beta^*) - \psi\lambda^*] + r^*\,[y(\beta_r^*) - \zeta\psi\lambda^*] - \omega\delta i^*$. Now use $g(i^*) = \phi - \mu - \delta i^*$ and equation (A.6) to obtain

$$
V^* = \frac{s^*\,[y(\beta^*) - \psi\lambda^*] + r^*\,[y(\beta_r^*) - \zeta\psi\lambda^*] - \omega\delta i^* + p\frac{y(\overline{\beta})}{\rho + \mu - \phi}}{\sigma - \phi + \delta i^*},
\tag{A.14}
$$

where $\frac{y(\overline{\beta})}{\rho + \mu - \phi} = \widetilde{V}$. Add ω to both sides of equation (A.14) to obtain

$$
V^* + \omega = \frac{s^*\,[y(\beta^*) - \psi\lambda^*] + r^*\,[y(\beta_r^*) - \zeta\psi\lambda^*] + p\frac{y(\overline{\beta})}{\rho + \mu - \phi}}{\sigma - \phi + \delta i^*} + \frac{\sigma - \phi}{\sigma - \phi + \delta i^*}\omega.
\tag{A.15}
$$

44

Notably, the coefficient on ω in (A.15) is positive, since $\sigma - \phi = \rho - (\phi - \mu) + p$ and $\rho > \phi - \mu$ from equation (27).

Inspection of (A.10) reveals that in the steady state, in which $\frac{dV_r}{dt} = 0$, $(\sigma + \chi) V_r^* = y(\beta_r^*) - \zeta\psi\lambda^* + \chi V_s^*$ so that

$$V_r^* = \frac{y(\beta_r^*) - \zeta\psi\lambda^* + \chi V_s^*}{\sigma + \chi}. \tag{A.16}$$

Equation (A.16) allows us to write the 3x3 system in V_s^*, V_i^* and V_r^* in equation (A.11) as the following 2x2 system in V_s^* and V_i^*

$$M^* \begin{bmatrix} V_s^* \\ V_i^* \end{bmatrix} = \begin{bmatrix} y(\beta^*) - \psi\lambda^* + \lambda^* \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} \\ (\gamma + \delta r^*) \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} - \delta(V^* + \omega) \end{bmatrix}. \tag{A.17}$$

where

$$M^* \equiv \begin{bmatrix} \sigma + \beta^* i^* + \lambda^* \frac{\sigma}{\sigma+\chi} & -\beta^* i^* \\ (\beta - \delta) s^* - (\gamma + \delta r^*)\frac{\chi}{\sigma+\chi} & \sigma + \delta + \gamma - \delta i^* - \beta^* s^* \end{bmatrix}. \tag{A.18}$$

Since $0 \le \lambda \le \lambda_{\max}$ and ω are both finite and $\rho - \phi + \mu + \delta i^* \ge \rho - \phi + \mu > 0$, equation (A.15) implies that $V^* + \omega$ is finite. Therefore, the vector on the right hand side of equation (A.17) is finite, which together with $\det M^* \ne 0$ (proved in Lemma 7) proves that V_s^* and V_i^* are finite. Since V_s^* is finite, equation (A.16) implies that V_r^* is finite.

The 2 x 2 system in equation (A.17) can be rewritten as

$$\begin{bmatrix} V_s^* \\ V_i^* \end{bmatrix} = M^{*-1} \begin{bmatrix} y(\beta^*) - \psi\lambda^* + \lambda^* \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} \\ (\gamma + \delta r^*) \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} - \delta(V^* + \omega) \end{bmatrix} \tag{A.19}$$

where

$$M^{*-1} = \frac{1}{\det M^*} \begin{bmatrix} \sigma + \delta + \gamma - \delta i^* - \beta^* s^* & \beta^* i^* \\ -\left[(\beta - \delta) s^* - (\gamma + \delta r^*)\frac{\chi}{\sigma+\chi}\right] & \sigma + \beta^* i^* + \lambda^* \frac{\sigma}{\sigma+\chi} \end{bmatrix}. \tag{A.20}$$

45

Substituting equation (A.20) into equation (A.19) and performing matrix multiplication yields

$$
\begin{bmatrix} V_s^* \\ V_i^* \end{bmatrix} = \frac{1}{\det M^*} \begin{bmatrix} \left(\begin{array}{c} (\sigma + \delta + \gamma - \delta i^* - \beta^* s^*) \left[y\left(\beta^*\right) - \psi\lambda^* + \lambda^* \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} \right] \\ + \beta^* i^* \left[(\gamma + \delta r^*) \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} - \delta\left(V^* + \omega\right) \right] \end{array} \right) \\ \left(\begin{array}{c} -\left[(\beta - \delta)s^* - (\gamma + \delta r^*)\frac{\chi}{\sigma+\chi} \right] \left[y\left(\beta^*\right) - \psi\lambda^* + \lambda^* \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} \right] \\ + \left(\sigma + \beta^* i^* + \lambda^* \frac{\sigma}{\sigma+\chi} \right) \left[(\gamma + \delta r^*) \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} - \delta\left(V^* + \omega\right) \right] \end{array} \right) \end{bmatrix}.
$$

(A.21)

Since $s^* + i^* + r^* = 1$, $\delta - \delta i^* - \delta s^* = \delta r^*$ so equation (A.21) implies

$$
V_s^* - V_i^* = \frac{1}{\det M^*} \left(\begin{array}{c} \left[\sigma + (\gamma + \delta r^*)\frac{\sigma}{\sigma+\chi} \right] \left[y\left(\beta^*\right) - \psi\lambda^* + \lambda^* \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} \right] \\ - \left(\sigma + \lambda^* \frac{\sigma}{\sigma+\chi} \right) \left[(\gamma + \delta r^*) \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} - \delta\left(V^* + \omega\right) \right] \end{array} \right),
$$

(A.22)

which can be rewritten as

$$
V_s^* - V_i^* = \frac{\sigma}{\det M^*} \left(\begin{array}{c} \left[1 + (\gamma + \delta r^*)\frac{1}{\sigma+\chi} \right] \left[y\left(\beta^*\right) - \psi\lambda^* + \lambda^* \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} \right] \\ - \left(1 + \lambda^* \frac{1}{\sigma+\chi} \right) \left[(\gamma + \delta r^*) \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} - \delta\left(V^* + \omega\right) \right] \end{array} \right)
$$

(A.23)

and further rearranged to obtain

$$
V_s^* - V_i^* = \frac{\sigma}{\det M^*} \left(\begin{array}{c} \left[1 + (\gamma + \delta r^*)\frac{1}{\sigma+\chi} \right] \left[y\left(\beta^*\right) - \psi\lambda^* \right] \\ - (\gamma + \delta r^* - \lambda^*) \frac{y(\beta_r^*) - \zeta\psi\lambda^*}{\sigma+\chi} + \left(1 + \lambda^* \frac{1}{\sigma+\chi} \right) \delta\left(V^* + \omega\right) \end{array} \right)
$$

(A.24)

and finally

$$
V_s^* - V_i^* = \frac{\sigma}{\det M^*} \left(\begin{array}{c} \left(1 + \frac{\lambda^*}{\sigma+\chi} \right) \left[y\left(\beta^*\right) - \psi\lambda^* + \delta\left(V^* + \omega\right) \right] \\ + \frac{\gamma + \delta r^* - \lambda^*}{\sigma+\chi} \left[y\left(\beta^*\right) - \psi\lambda^* - \left(y\left(\beta_r^*\right) - \zeta\psi\lambda^* \right) \right] \end{array} \right).
$$

(A.25)

A.1.1. Steady State Iso-infection Lines

An iso-infection locus is the set of pairs (β, λ) for which $i^* > 0$ is constant. Use the implicit function theorem along with Proposition 6, which implies that $\frac{di^*}{d\beta} = -\beta^{-1}\left(\eta^* + \lambda\right)\frac{di^*}{d\lambda}$, so

46

that

$$\frac{d\lambda}{d\beta}\Big|_{i^* \text{ constant}} = \frac{\eta^* + \lambda}{\beta}. \tag{A.26}$$

Now consider the iso-infection locus through point E in Figure 1, where $(\beta, \lambda) = (\beta_E, 0)$ for an arbitrary $\beta_E \in (\beta_c, \overline{\beta})$. Let η_E^* be the root of the quadratic equation $q(\eta) = 0$ when $\beta = \beta_E$ and $\lambda = 0$ in the quadratic function $q(\eta)$ defined in Lemma 4. (Note that $\eta^* = \eta_E^*$ because i^* stays constant). All points on the iso-infection locus through Point E satisfy the ODE in equation (A.26) subject to the boundary condition that $\lambda = 0$ when $\beta = \beta_E$. The linear equation

$$\lambda = \frac{\eta_E^*}{\beta_E} \beta - \eta_E^* \tag{A.27}$$

has slope $\frac{\eta_E^*}{\beta_E} = \frac{\eta_E^* + \lambda}{\beta}$ and thus satisfies the ODE in equation (A.26). It also satisfies the boundary condition that $\lambda = 0$ when $\beta = \beta_E$.

Finally, consider an iso-infection locus, which we can now call an iso-infection line, through Point F, which lies to right of Point E (and to the left of Point C) on the horizontal axis in Figure 1. Let β_F be the value of β at Point F. Since $\beta_F > \beta_E$ and $\lambda = 0$ at both points, the infection rate i_F^* on in the iso-infection line through Point F is higher than i_E^*, the infection rate on the iso-infection line through Point E, which implies that $\eta_F^* < \eta_E^*$. Therefore, $\frac{\eta_F^*}{\beta_F} < \frac{\eta_E^*}{\beta_E}$ so the slope of the iso-infection line through Point F is less than the slope of the iso-infection line through Point E.

B   Selected Proofs

Proof of Proposition 2. Assume that $i^* = 0$. Then equation (15) implies $\eta^* = \phi + \chi$, so equation (17) implies $\lambda s^* = \eta^* r^* = (\phi + \chi) r^*$. Since $s^* + r^* = 1$, it follows that $s^* = 1 - \frac{\lambda}{\phi + \chi + \lambda}$ and $r^* = \frac{\lambda}{\phi + \chi + \lambda}$.   ∎

Proof of Proposition 3. Assume that $i^* > 0$. Then equation (16) implies $\beta s^* =$

$(\beta_c - \delta i^*)$, equivalently, $s^* = \frac{\beta_c - \delta i^*}{\beta}$. From the definition of $\mathcal{R}_0^*$ in equation (11), $\mathcal{R}_0^{*-1} = \frac{\delta + \gamma + g^* + \mu}{\beta} = \frac{\delta + \gamma + \phi - \delta i^*}{\beta} = \frac{\beta_c - \delta i^*}{\beta}$, which completes the proof of Statement 1. Substitute $\mathcal{R}_0^{*-1}$ for s^* in equation (17) to obtain $\gamma i^* + \lambda \mathcal{R}_0^{*-1} = \eta^* r^*$, which together with $i^* + r^* = 1 - \mathcal{R}_0^{*-1}$, are two linear equations in i^* and r^*. The solution of this 2-equation linear system is given by Statements 2 and 3. ∎

Lemma 4 *Consider the quadratic function in equation* (19)*. Then*

1. *$q(0) = -\left[\beta\gamma(\phi + \chi) + \delta\lambda(\delta + \gamma - \chi)\right] < 0$.*

2. *$q(\phi + \chi) = (\phi + \chi)\delta\left[\beta - \left(1 + \frac{\lambda}{\phi + \chi}\right)\beta_c\right]$*

3. *If $i^* > 0$, then*

   - *$\beta > \delta$ so that $q(\eta)$ is convex*

   - *equation (15) implies that $\eta^* \equiv \mu + g^* + \chi = \phi - \delta i^* + \chi < \phi + \chi$ is the positive root of $q(\eta) = 0$.*

Proof of Lemma 4. (Statement 1) Evaluate $q(\eta)$ at $\eta = 0$ to obtain $q(0) = -(\beta\gamma - \delta\lambda)(\phi + \chi) - \delta\lambda\beta_c = -\left[(\beta\gamma - \delta\lambda)(\phi + \chi) + \delta\lambda(\delta + \gamma + \phi)\right] = -\left[\beta\gamma(\phi + \chi) + \delta\lambda(\delta + \gamma - \chi)\right] < 0$, where the inequality follows from equation (4). (Statement 2) Evaluate $q(\eta)$ at $\eta = \phi + \chi$ to obtain $q(\phi + \chi) = \eta\delta\left[\beta - \left(1 + \frac{\lambda}{\eta}\right)\beta_c\right] = (\phi + \chi)\delta\left[\beta - \left(1 + \frac{\lambda}{\phi + \chi}\right)\beta_c\right]$.

(Statement 3) Assume that $i^* > 0$. Therefore, $\beta s^* = \delta + \gamma + \mu + g^* = \delta + \gamma + \phi - \delta i^*$, where the first equality follows from equation (7) and the second equality uses $\mu + g^* = \phi - \delta i^*$ from equation (15). Therefore, $(\beta - \delta)s^* = \delta + \gamma + \phi - \delta i^* - \delta s^* = \gamma + \phi + \delta r^* > 0$, which implies $\beta - \delta > 0$. Hence, the second-order term in $q(\eta)$ is $(\beta - \delta)\eta^2 > 0$ so $q(\eta)$ is convex. Therefore, since $q(0) < 0$, $q(\eta) = 0$ has a unique positive root. Also since $i^* > 0$, Statement 1 of Proposition 3 implies $s^* = \frac{\beta_c - \delta i^*}{\beta}$. In equation (17), replace r^* by $(1 - s^* - i^*)$ to obtain $(\gamma + \eta^*)i^* + (\lambda + \eta^*)s^* = \eta^*$. Now replace i^* by $\frac{\phi + \chi - \eta^*}{\delta}$, replace s^* by $\frac{\beta_c - (\phi + \chi - \eta^*)}{\beta}$ and multiply both sides of the resulting equation by $\beta\delta$ to obtain $\beta(\gamma + \eta^*)(\phi + \chi - \eta^*) + \delta(\lambda + \eta^*)(\beta_c - (\phi + \chi - \eta^*)) = \beta\delta\eta^*$, which can be rewritten as

$[(\beta - \delta) \eta^* + \beta\gamma - \delta\lambda] (\phi + \chi - \eta^*) - \delta [\beta\eta^* - (\lambda + \eta^*) \beta_c] = 0$. Finally, multiply both sides by -1 to obtain $q(\eta^*) \equiv [(\beta - \delta) \eta^* + \beta\gamma - \delta\lambda] (\eta^* - (\phi + \chi)) + \delta [\beta\eta^* - (\lambda + \eta^*) \beta_c] = 0$. ∎

Proof of Proposition 4. (i) Assume that $\beta < \left(1 + \frac{\lambda}{\phi + \chi}\right) \beta_c$ and, contrary to what is to be proved, $i^* > 0$. Since $i^* > 0$, Lemma 4 implies that η^* is the unique root of $q(\eta) = 0$ in $(0, \phi + \chi]$. However, Lemma 4 also implies that $q(\eta)$ is convex, $q(0) < 0$, and $q(\phi + \chi) < 0$, so $q(\eta) = 0$ does not have any roots in $(0, \phi + \chi]$. Therefore, $i^* = 0$ and $\mathcal{R}_0^* = \frac{\beta}{\beta_c - \delta i^*} = \frac{\beta}{\beta_c} < 1 + \frac{\lambda}{\phi + \chi}$. (ii) Assume that $\beta = \left(1 + \frac{\lambda}{\phi + \chi}\right) \beta_c$. Lemma 4 implies that $q(\phi + \chi) = 0$, so $\eta^* = \phi + \chi$ and $\delta i^* = \phi + \chi - \eta^* = 0$, so $i^* = 0$ and $\mathcal{R}_0^* = \frac{\beta}{\beta_c - \delta i^*} = \frac{\beta}{\beta_c} = 1 + \frac{\lambda}{\phi + \chi}$. (iii) Assume that $\beta > \left(1 + \frac{\lambda}{\phi + \chi}\right) \beta_c$. Suppose, contrary to what is to be shown, that $i^* = 0$. Proposition 2 implies $s^* = 1 - \frac{\lambda}{\phi + \chi + \lambda} = \frac{\phi + \chi}{\phi + \chi + \lambda}$, so $\beta s^* > \left(1 + \frac{\lambda}{\phi + \chi}\right) \beta_c \frac{\phi + \chi}{\phi + \chi + \lambda} = \beta_c$ and hence $\frac{di}{dt} > 0$ in the positive neighborhood of $i^* = 0$. Therefore, i will not approach zero, so $i^* > 0$. Since $\beta > \left(1 + \frac{\lambda}{\phi + \chi}\right) \beta_c$, Lemma 4 implies that $q(\phi + \chi) > 0$, so $q(\eta) = 0$ has a unique root in $(0, \phi + \chi]$, which is $\eta^* = \phi + \chi - \delta i^*$, so $i^* > 0$. Also, $\mathcal{R}_0^* = \frac{\beta}{\beta_c - \delta i^*} > \frac{\beta}{\beta_c} > 1 + \frac{\lambda}{\phi + \chi}$. ∎

Proof of Proposition 5. Statement 1: Consider a particular $\lambda_0 \geq 0$ and $\beta_0 \equiv \left(1 + \frac{\lambda_0}{\phi + \chi}\right) \beta_c$. Proposition 4 implies that if $\beta = \beta_0$ and $\lambda = \lambda_0$, then $i^* = 0$. It also implies that if $\beta = \beta_0$ and $\lambda < \lambda_0$, then $\beta > \left(1 + \frac{\lambda}{\phi + \chi}\right) \beta_c$ and $i^* > 0$. Therefore, λ_0 is the minimum value of λ for which $i^* = 0$ when $\beta = \beta_0$. Rearranging the definition of $\beta_0 \equiv \left(1 + \frac{\lambda_0}{\phi + \chi}\right) \beta_c$ as $\lambda_0 \equiv \left(\frac{\beta_0}{\beta_c} - 1\right) (\phi + \lambda)$ thus implies that $\lambda_h(\beta) = \left(\frac{\beta}{\beta_c} - 1\right) (\phi + \chi)$. Statement 2: If $\lambda = \lambda_h(\beta)$, then $i^* = 0$ so $g^* = \phi - \mu$, so $\mathcal{R}_0 \equiv \frac{\beta}{\delta + \gamma + g + \mu} = \frac{\beta}{\delta + \gamma + \phi} = \frac{\beta}{\beta_c}$. Observe that $1 + \frac{\lambda}{\phi + \chi} = 1 + \frac{\lambda_h(\beta)}{\phi + \chi} = 1 + \left(\frac{\beta}{\beta_c} - 1\right) = \frac{\beta}{\beta_c}$. Therefore, $\mathcal{R}_0 = 1 + \frac{\lambda}{\phi + \chi}$, so Proposition 4 implies $s^* = \mathcal{R}_0^{*-1}$ and $r^* = \frac{\gamma}{\eta^* + \gamma} \left(1 - \mathcal{R}_0^{*-1}\right) + \frac{\lambda}{\eta^* + \gamma} \mathcal{R}_0^{*-1}$. Since $\mathcal{R}_0 = \frac{\beta}{\beta_c}$, we have $s^* = \frac{\beta_c}{\beta}$ and $r^* = \frac{\gamma}{\eta^* + \gamma} \left(1 - \frac{\beta_c}{\beta}\right) + \frac{\lambda}{\eta^* + \gamma} \frac{\beta_c}{\beta}$. Since $i^* = 0$, equation (15) implies that $\eta^* = \phi + \chi$, so $r^* = \frac{\gamma}{\phi + \chi + \gamma} \left(1 - \frac{\beta_c}{\beta}\right) + \frac{\lambda}{\phi + \chi + \gamma} \frac{\beta_c}{\beta}$ and since $\lambda_h(\beta) = \left(\frac{\beta}{\beta_c} - 1\right) (\phi + \chi)$, we have $r^* = \frac{\gamma}{\phi + \chi + \gamma} \left(1 - \frac{\beta_c}{\beta}\right) + \frac{\phi + \chi}{\phi + \chi + \gamma} \left(\frac{\beta}{\beta_c} - 1\right) \frac{\beta_c}{\beta} = \frac{\gamma}{\phi + \chi + \gamma} \left(1 - \frac{\beta_c}{\beta}\right) + \frac{\phi + \chi}{\phi + \chi + \gamma} \left(1 - \frac{\beta_c}{\beta}\right) = 1 - \frac{\beta_c}{\beta}$. ∎

Proof of Proposition 6. Since, $\beta > \widetilde{\beta}_c(\lambda)$, $i^* > 0$ so Lemma 4 Statement 3 implies that $q(\eta) \equiv [(\beta - \delta) \eta + \beta\gamma - \delta\lambda] (\eta - (\phi + \chi)) + \delta [\beta\eta - (\eta + \lambda) \beta_c]$ is convex. Since $\eta^* \geq 0$, implies that it is the larger root of $q(\eta^*) = 0$, and hence $q'(\eta^*) > 0$. Since $q(\eta)$ is linear in β, it can be written as $q(\eta) = \beta \frac{dq}{d\beta} - \delta [(\eta + \lambda) (\eta - (\phi + \chi)) + (\eta + \lambda) \beta_c]$. Use $\eta \equiv \phi + \chi - \delta i$

and $\beta_c \equiv \delta + \gamma + \phi$ to obtain $q(\eta) = \beta \frac{dq}{d\beta} - \delta(\eta + \lambda)[(1-i)\delta + \gamma + \phi]$. Therefore, $0 = q(\eta^*) = \beta \frac{dq(\eta^*)}{d\beta} - \delta(\eta^* + \lambda)[(1-i)\delta + \gamma + \phi]$, which implies $\frac{dq(\eta^*)}{d\beta} = \beta^{-1}\delta(\eta^* + \lambda)[(1-i)\delta + \gamma + \phi] >$ 0. Therefore, $\frac{d\eta^*}{d\beta} = -\frac{dq(\eta^*)}{d\beta}[q'(\eta^*)]^{-1} < 0$. Since $\frac{di^*}{d\eta^*} = -\frac{1}{\delta}$, $\frac{di^*}{d\beta} = \beta^{-1}(\eta^* + \lambda)[(1-i)\delta + \gamma + \phi][q'(\eta^*)]^{-1}$ 0. Differentiate $q(\eta)$ with respect to λ to obtain $\frac{dq}{d\lambda} = -\delta[(\eta^* - (\phi + \chi)) + \beta_c] = -\delta[(1-i)\delta + \gamma + \phi] <$ 0. $\frac{d\eta^*}{d\lambda} = -\frac{dq(\eta^*)}{d\lambda}[q'(\eta^*)]^{-1} = \delta[(1-i)\delta + \gamma + \phi][q'(\eta^*)]^{-1} > 0$. Since $\frac{di^*}{d\eta^*} = -\frac{1}{\delta}$, $\frac{di^*}{d\lambda} = -\frac{1}{\delta}\frac{d\eta^*}{d\lambda} = -[(1-i)\delta + \gamma + \phi][q'(\eta^*)]^{-1} < 0.$ ∎

Proof of Lemma 1. Lemma 1 is a corollary of Lemma 7 with $\lambda^* = 0$. ∎

Proof of Proposition 8. In the steady state, $\frac{ds}{dt} = \frac{di}{dt} = \frac{dr}{dt} = 0$ and using $\sigma \equiv \rho + \mu + p$ and equation (35) to substitute $\frac{y(\overline{\beta}) + \chi v^S}{(\rho + \mu + p + \chi)}$ for v^R, the three-equation system in equations (35)-(37) can be written as the following two-equation system

$$\sigma v^S = y(\beta_{LF}^*) + \beta_{LF}^* i_{LF}^* (v^I - v^S) \tag{B.1}$$

and

$$(\sigma + \delta) v^I = -\omega\delta + \gamma\left(\frac{y(\overline{\beta}) + \chi v^S}{\sigma + \chi} - v^I\right). \tag{B.2}$$

This two-equation system can be written as

$$M_{LF}^* \begin{bmatrix} v^S \\ v^I \end{bmatrix} = \begin{bmatrix} y(\beta_{LF}^*) \\ \gamma\frac{y(\overline{\beta})}{\sigma + \chi} - \omega\delta \end{bmatrix} \tag{B.3}$$

where

$$M_{LF}^* = \begin{bmatrix} \sigma + \beta_{LF}^* i_{LF}^* & -\beta_{LF}^* i_{LF}^* \\ -\gamma\frac{\chi}{\sigma + \chi} & \sigma + \delta + \gamma \end{bmatrix}. \tag{B.4}$$

Premultiply both sides of equation (B.3) by

$$M_{LF}^{*-1} = \frac{1}{\det M_{LF}^*}\begin{bmatrix} \sigma + \delta + \gamma & \beta_{LF}^* i_{LF}^* \\ \gamma\frac{\chi}{\sigma + \chi} & \sigma + \beta_{LF}^* i_{LF}^* \end{bmatrix} \tag{B.5}$$

50

to obtain

$$\begin{bmatrix} v^S \\ v^I \end{bmatrix} = \frac{1}{\det M_{LF}^*} \begin{bmatrix} (\sigma + \delta + \gamma) \, y \, (\beta_{LF}^*) + \beta_{LF}^* i_{LF}^* \left(\gamma \frac{y(\bar{\beta})}{\sigma + \chi} - \omega \delta \right) \\ \gamma \frac{\chi}{\sigma + \chi} y \, (\beta_{LF}^*) + (\sigma + \beta_{LF}^* i_{LF}^*) \left(\gamma \frac{y(\bar{\beta})}{\sigma + \chi} - \omega \delta \right) \end{bmatrix}. \tag{B.6}$$

Subtract the second row from the first row on each side of equation (B.6) and rearrange to obtain

$$v^S - v^I = \frac{1}{\det M_{LF}^*} \left[(\sigma + \delta) \, y \, (\beta_{LF}^*) + \gamma \frac{\sigma}{\sigma + \chi} \left(y \, (\beta_{LF}^*) - y \, (\bar{\beta}) \right) + \sigma \omega \delta \right]. \tag{B.7}$$

Define

$$h \, (\beta) \equiv (\sigma + \delta) \, y \, (\beta) + \gamma \frac{\sigma}{\sigma + \chi} \left(y \, (\beta) - y \, (\bar{\beta}) \right) + \sigma \omega \delta \tag{B.8}$$

$$d \, (\beta) \equiv \sigma \, (\sigma + \delta + \gamma) + \beta i^* \, (\beta) \left(\sigma + \delta + \gamma \frac{\sigma}{\sigma + \chi} \right) > 0 \tag{B.9}$$

and

$$w \, (\beta) \equiv \frac{h \, (\beta) \, i^* \, (\beta)}{d \, (\beta)} \tag{B.10}$$

so that $h \, (\beta_{LF}^*)$ is the expression in square brackets in equation (B.7) and $d \, (\beta_{LF}^*)$ is equal to $\det M_{LF}^*$. Accordingly,

$$w \, (\beta_{LF}^*) = \left(v^S - v^I \right) i^* \, (\beta_{LF}^*) \tag{B.11}$$

and the first-order condition in equation (38) can be written as

$$y' \, (\beta_{LF}^*) \leq w \, (\beta_{LF}^*) = \frac{h \, (\beta_{LF}^*) \, i^* \, (\beta_{LF}^*)}{d \, (\beta)}, \tag{B.12}$$

51

where the inequality reflects the constraint $i \geq 0$.

Lemma 5 $w'\left(\beta_{LF}^*\right) = \sigma\left(\sigma + \delta + \gamma\right) \dfrac{h\left(\beta_{LF}^*\right)}{\left[d\left(\beta_{LF}^*\right)\right]^2} i^{*\prime}\left(\beta_{LF}^*\right) > 0.$

Proof of of Lemma 5. First, we prove that $h\left(\beta_{LF}^*\right) > 0$. We consider two cases: (i) If $\beta_{LF}^* = \overline{\beta}$, use equation (B.8) to obtain $h\left(\beta_{LF}^*\right) = h\left(\overline{\beta}\right) = \left(\sigma + \delta\right) y\left(\overline{\beta}\right) + \sigma\omega\delta > 0$. (ii) If $\beta_{LF}^* < \overline{\beta}$, there are two sub-cases: (a) if the first-order condition in equation (B.12) holds as an equality, then $0 < y'\left(\beta_{LF}^*\right) = \dfrac{h\left(\beta_{LF}^*\right) i^*\left(\beta_{LF}^*\right)}{d(\beta)}$, and since equation (B.9) states that $d\left(\beta_{LF}^*\right) > 0$, equation (B.12) implies that $h\left(\beta_{LF}^*\right) i^*\left(\beta_{LF}^*\right) > 0$, which implies $h\left(\beta_{LF}^*\right) > 0$. (b) if the first-order condition in equation (B.12) holds as a strict inequality, then $i^*\left(\beta_{LF}^*\right) = 0$, which, along with $y'\left(\beta_{LF}^*\right) > 0$ violates equation (B.12). Therefore, we need not consider this sub-case.

Differentiate $h\left(\beta\right)$ and $d\left(\beta\right)$ in equations (B.8) and (B.9), respectively, and evaluate these derivatives at $\beta = \beta_{LF}^*$, using the first-order condition in equation (B.12) to obtain

$$h'\left(\beta_{LF}^*\right) = \left(\sigma + \delta + \gamma\frac{\sigma}{\sigma+\chi}\right) y'\left(\beta_{LF}^*\right) = \left(\sigma + \delta + \gamma\frac{\sigma}{\sigma+\chi}\right) \frac{h\left(\beta_{LF}^*\right) i^*\left(\beta_{LF}^*\right)}{d\left(\beta_{LF}^*\right)} \quad \text{(B.13)}$$

and

$$d'\left(\beta_{LF}^*\right) = \left(\sigma + \delta + \gamma\frac{\sigma}{\sigma+\chi}\right) \left(i^*\left(\beta_{LF}^*\right) + \beta_{LF}^* i^{*\prime}\left(\beta_{LF}^*\right)\right). \quad \text{(B.14)}$$

Differentiate $w\left(\beta\right)$ in equation (B.10) and evaluate the derivative at $\beta = \beta_{LF}^*$ to obtain

$$w'\left(\beta_{LF}^*\right) = \frac{\left[h'\left(\beta_{LF}^*\right) d\left(\beta_{LF}^*\right) - h\left(\beta_{LF}^*\right) d'\left(\beta_{LF}^*\right)\right] i^*\left(\beta_{LF}^*\right) + h\left(\beta_{LF}^*\right) i^{*\prime}\left(\beta_{LF}^*\right) d\left(\beta_{LF}^*\right)}{\left[d\left(\beta_{LF}^*\right)\right]^2}. \quad \text{(B.15)}$$

Use equations equations (B.8), (B.9), and (B.13) and rearrange to calculate

$$h'\left(\beta_{LF}^*\right) d\left(\beta_{LF}^*\right) - h\left(\beta_{LF}^*\right) d'\left(\beta_{LF}^*\right) = -h\left(\beta_{LF}^*\right) \left(\sigma + \delta + \gamma\frac{\sigma}{\sigma+\chi}\right) \beta_{LF}^* i^{*\prime}\left(\beta_{LF}^*\right) \quad \text{(B.16)}$$

Substitute equation (B.16) into equation (B.15) to obtain

$$w'\left(\beta_{LF}^*\right) = \sigma\left(\sigma + \delta + \gamma\right) \frac{h\left(\beta_{LF}^*\right)}{\left[d\left(\beta_{LF}^*\right)\right]^2} i^{*\prime}\left(\beta_{LF}^*\right) > 0. \quad \text{(B.17)}$$

52

When it holds as an equality, the first-order condition in the laissez-faire case in equation (B.12) can be interpreted as equating the marginal benefit of increasing β, which is $y'(\beta)$, and the marginal cost of increasing β, which is $w(\beta)$. The marginal benefit, $y'(\beta)$, is decreasing in β and when $y'(\beta) = w(\beta)$, Lemma 5 implies that the marginal cost, $w(\beta)$, is increasing in β. Therefore, there is a unique value of β, β_{LF}^*, where $y'(\beta) = w(\beta)$. In addition, equation (B.8) implies that $h(\overline{\beta}) > 0$ so that along with $i^*(\overline{\beta}) > 0$ and $d(\overline{\beta}) > 0$, $w(\overline{\beta}) > 0 = y'(\overline{\beta})$. Therefore, $\beta_{LF}^* < \overline{\beta}$. ∎

Proof of Proposition 9. Set $\delta = 0$ (as well as $\lambda = 0$) in $V_s^* - V_i^*$ in equation (A.25) and since $i^*(\beta_{OP}^*) > 0$, use $\beta s^* = \delta + \gamma + \phi - \delta i^* = \gamma + \phi$ to obtain

$$(V_s^* - V_i^*) i^*(\beta_{OP}^*) = W(\beta_{OP}^*), \tag{B.18}$$

where

$$W(\beta) \equiv \frac{H(\beta) i^*(\beta)}{D(\beta)}, \tag{B.19}$$

$$H(\beta) \equiv \sigma y(\beta) + \gamma \frac{\sigma}{\sigma + \chi} \left(y(\beta) - y(\overline{\beta}) \right), \qquad \text{since } \delta = 0 \tag{B.20}$$

and

$$D(\beta) \equiv \sigma(\sigma - \phi) + \beta i^*(\beta) \left(\sigma + \gamma \frac{\sigma}{\sigma + \chi} \right), \qquad \text{since } \delta = 0. \tag{B.21}$$

Set $\delta = 0$ in equation (B.8) to obtain

$$h(\beta) = \sigma y(\beta) + \gamma \frac{\sigma}{\sigma + \chi} \left(y(\beta) - y(\overline{\beta}) \right), \qquad \text{since } \delta = 0 \tag{B.22}$$

and observe that $H(\beta)$ and $h(\beta)$ are identically equal when $\delta = 0$. Now set $\delta = 0$ in equation

(B.9) to obtain

$$d\left(\beta\right) = \sigma\left(\sigma + \gamma\right) + \beta i^{*}\left(\beta\right)\left(\sigma + \gamma \frac{\sigma}{\sigma + \chi}\right), \qquad \text{since } \delta = 0. \tag{B.23}$$

Subtracting equation (B.21) from equation (B.23) yields

$$d\left(\beta\right) - D\left(\beta\right) = \sigma\left(\gamma + \phi\right) > 0, \qquad \text{since } \delta = 0. \tag{B.24}$$

Therefore, when $\delta = 0$, $W\left(\beta\right) \equiv \frac{H(\beta)i^{*}(\beta)}{D(\beta)} > w\left(\beta\right) \equiv \frac{h(\beta)i^{*}(\beta)}{d(\beta)}$ because $H\left(\beta\right)$ and $h\left(\beta\right)$ are identically equal and $d\left(\beta\right) > D\left(\beta\right)$. Lemma 5 implies that there is a unique β_{LF}^{*} for which $y'\left(\beta_{LF}^{*}\right) = w\left(\beta_{LF}^{*}\right)$ and $y'\left(\beta\right) < w\left(\beta\right)$ for $\beta > \beta_{LF}^{*}$. Since $W\left(\beta\right) > w\left(\beta\right)$, we have $W\left(\beta\right) > w\left(\beta\right) \geq y'\left(\beta\right)$ for $\beta \geq \beta_{LF}^{*}$. Therefore, $W\left(\beta_{OP}^{*}\right) = y'\left(\beta_{OP}^{*}\right)$ implies $\beta_{OP}^{*} < \beta_{LF}^{*}$. ∎

Proof of Lemma 3. At all points on $\overline{BD}$, $i^{*} = 0$, $s^{*} + r^{*} = 1$, and $\beta = \overline{\beta}$, so the instantaneous steady-state flow of aggregate output, $s^{*}y\left(\beta\right) + r^{*}y\left(\overline{\beta}\right)$, is $y\left(\overline{\beta}\right)$. The per capita aggregate cost of the vaccine is $\Psi = \psi\lambda\left(s + \zeta r\right)$. In the steady state, this cost is Ψ^{*} $= \psi\lambda^{*}\left(s^{*} + \zeta r^{*}\right) = \psi\lambda^{*}\left(\left(1 - \zeta\right)s^{*} + \zeta\right)$. Since $s^{*} = 1 - \frac{\lambda^{*}}{\phi + \chi + \lambda^{*}}$ at all points on $\overline{BD}$, we have $s^{*} = \frac{\phi + \chi}{\phi + \chi + \lambda^{*}} > 0$ and $\frac{ds^{*}}{d\lambda^{*}} = -\frac{s^{*}}{\phi + \chi + \lambda^{*}} < 0$ so $\frac{d\Psi^{*}}{d\lambda^{*}} = \psi\left(\left(1 - \zeta\right)s^{*} + \zeta\right) - \psi\lambda^{*}\left(1 - \zeta\right)\frac{s^{*}}{\phi + \chi + \lambda^{*}} = \psi\left[\left(1 - \zeta\right)s^{*}\frac{\phi + \chi}{\phi + \chi + \lambda^{*}} + \zeta\right] > 0$. Therefore, Ψ^{*} is minimized along $\overline{BD}$ at point B and the instantaneous steady-state flow of social welfare, $y\left(\overline{\beta}\right) - \Psi^{*}$, is maximized along $\overline{BD}$ at Point B. To go beyond this comparative statics result, consider an arbitrary Point J that lies on $\overline{BD}$ above Point B. Start from the steady state at Point J and consider an immediate and permanent reduction in λ to $\lambda_{h}\left(\overline{\beta}\right)$, which is the value of λ at Point B. Along the transition path to the new steady state at Point B, $i = 0$, which implies $g = \phi - \mu$ and $s + r = 1$, so the dynamic behavior of s in equation (6) becomes $\frac{ds}{dt} = \phi + \chi r - \left(\phi + \lambda\right)s = \phi + \chi - \left(\phi + \chi + \lambda\right)s$, which is a first-order linear constant coefficient ODE. Starting from Point J, the permanent reduction in λ to $\lambda_{h}\left(\overline{\beta}\right)$ causes $\frac{ds}{dt}$ to become positive so that s increases monotonically toward $1 - \frac{\lambda_{h}\left(\overline{\beta}\right)}{\phi + \chi + \lambda_{h}\left(\overline{\beta}\right)}$. Along the transition path, the permanent reduction in λ causes an immediate reduction in vaccination cost by $\left(\lambda_{J}^{*} - \lambda_{h}\left(\overline{\beta}\right)\right)\psi\left(\left(1 - \zeta\right)s_{J}^{*} + \zeta\right)$, where λ_{J}^{*} and s_{J}^{*} are the values of λ and s in the initial steady state at Point J. Along the transition path, the vaccination cost $\psi\lambda_{h}\left(\overline{\beta}\right)\left(\left(1 - \zeta\right)s + \zeta\right)$ increases monotonically toward

54

$\Psi^* = \psi \lambda_h \left(\overline{\beta} \right) \left((1 - \zeta) s_B^* + \zeta \right)$, where s_B^* is the value of s in the new steady state at Point B. Since we have already shown that Ψ^* is lower at Point B than at Point J, it follows that the permanent reduction in λ reduces the vaccination cost everywhere along the transition path to the new steady state, without changing aggregate output per capita or deaths from the disease. Therefore, Point J cannot be an optimal steady state. ∎

Lemma 6 *In the presence of a vaccine with $\psi > 0$, if $y\left(\overline{\beta}\right) > \psi \max \left\{ \sigma + \chi + \gamma + \delta, \left(\frac{\overline{\beta}}{\beta_c} - 1 \right) (\phi + \chi) \right\}$ and $\beta^* = \overline{\beta}$, then $sign\left(V_s^* - V_i^* \right) = sign\left(\det M^* \right)$.*

Proof of Lemma 6. We have shown that the optimal pair $(\beta_{OP}^*, \lambda_{OP}^*)$ must lie in or on the triangle ABC in Figure 1, so $\lambda_{OP}^* \leq \lambda_h\left(\overline{\beta}\right) = \left(\frac{\overline{\beta}}{\beta_c} - 1 \right)(\phi + \chi) < \frac{y(\overline{\beta})}{\psi}$, where the final inequality follows from the assumption $\frac{y(\overline{\beta})}{\psi} > \max \left\{ \sigma + \chi + \gamma + \delta, \left(\frac{\overline{\beta}}{\beta_c} - 1 \right)(\phi + \chi) \right\}$. Equation (A.15) implies that when $\beta^* = \overline{\beta}$, $V^* + \omega > 0$ if $\frac{y(\overline{\beta})}{\psi} > \lambda^*$, which implies $\frac{y(\overline{\beta})}{\psi} > \zeta \lambda^*$. Since $\zeta \in \{0, 1\}$, the statement above that $\lambda_{OP}^* < \frac{y(\overline{\beta})}{\psi}$ implies that $V^* + \omega > 0$ when $\beta^* = \overline{\beta}$. When $\beta^* = \overline{\beta}$, equation (A.25) implies that $V_s^* - V_i^* = \frac{\sigma}{\det M^*} \left(\begin{array}{c} \left(1 + \frac{\lambda^*}{\rho + \mu + \chi} \right) \left[y\left(\overline{\beta} \right) - \psi \lambda^* + \delta \left(V^* + \omega \right) \right] \\ - \frac{\gamma + \delta r^* - \lambda^*}{\sigma + \chi} (1 - \zeta) \psi \lambda^* \end{array} \right)$ and since $\delta\left(V^* + \omega \right) > 0$, $V_s^* - V_i^*$ will have the same sign as $\det M^*$ if $A \equiv \left(1 + \frac{\lambda^*}{\sigma + \chi} \right) \left[y\left(\overline{\beta} \right) - \psi \lambda^* \right] - \frac{\gamma + \delta r^* - \lambda^*}{\sigma + \chi} (1 - \zeta) \psi \lambda^*$ is positive. We consider two cases. **Case I**: $\zeta = 1$. When $\zeta = 1$, $A = \left(1 + \frac{\lambda^*}{\sigma + \chi} \right) \left[y\left(\overline{\beta} \right) - \psi \lambda^* \right] > 0$, where the inequality follows from $\lambda_{OP}^* < \frac{y(\overline{\beta})}{\psi}$. **Case II**: $\zeta = 0$. When $\zeta = 0$, $A \equiv \left(1 + \frac{\lambda^*}{\sigma + \chi} \right) \left[y\left(\overline{\beta} \right) - \psi \lambda^* \right] - \frac{\gamma + \delta r^* - \lambda^*}{\sigma + \chi} \psi \lambda^*$ will be positive if $(\sigma + \chi + \lambda^*) \left[y\left(\overline{\beta} \right) - \psi \lambda^* \right] > (\gamma + \delta r^* - \lambda^*) \psi \lambda^*$, which is equivalent to $(\sigma + \chi + \lambda^*) y\left(\overline{\beta} \right) > (\gamma + \delta r^* + \sigma + \chi) \psi \lambda^*$, which is equivalent to $\frac{y(\overline{\beta})}{\psi} > (\gamma + \delta r^* + \sigma + \chi) \frac{\lambda^*}{\sigma + \chi + \lambda^*}$. This inequality holds because $0 \leq \frac{\lambda^*}{\sigma + \chi + \lambda^*} < 1$ and the assumption that $\frac{y(\overline{\beta})}{\psi} > \max \left\{ \sigma + \chi + \gamma + \delta, \left(\frac{\overline{\beta}}{\beta_c} - 1 \right)(\phi + \chi) \right\}$. ∎

Corollary 1 (to Lemma 6) *In the absence of a vaccine, if $\beta^* = \overline{\beta}$, then $sign\left(V_s^* - V_i^* \right) = sign\left(\det M^* \right)$.*

Proof of Corollary 1. It suffices to prove that $A \equiv \left(1 + \frac{\lambda^*}{\sigma + \chi} \right) \left[y\left(\overline{\beta} \right) - \psi \lambda^* \right] - \frac{\gamma + \delta r^* - \lambda^*}{\sigma + \chi} (1 - \zeta) \psi \lambda^*$ in the proof of Lemma 6 is positive. Setting $\lambda^* = 0$ in the definition of A yields $A = y\left(\overline{\beta} \right) > 0$. ∎

Lemma 7 *Under optimal policy,* $\det M^* > 0$ *and* $V_s^* - V_i^* > 0$ *is finite.*

Proof of Lemma 7. Inspection of equation (A.25) reveals that $V_s^* - V_i^*$ is finite if $\det M^* > 0$. M^* is given in equation (A.18), which we repeat here for convenience,

$$
M^* \equiv \begin{bmatrix} \sigma + \beta^* i^* + \lambda^* \frac{\sigma}{\sigma+\chi} & -\beta^* i^* \\ (\beta - \delta) s^* - (\gamma + \delta r^*) \frac{\chi}{\sigma+\chi} & \sigma + \delta + \gamma - \delta i^* - \beta^* s^* \end{bmatrix} \tag{B.25}
$$

Case I: Assume $i^* > 0$. Setting $\frac{di}{dt} = 0$ in equation (A.2) implies that in the steady state $\beta^* s^* = \delta + \gamma + \phi - \delta i^*$. Therefore, $(\beta - \delta) s^* - (\gamma + \delta r^*) \frac{\chi}{\sigma+\chi} = \delta + \gamma + \phi - \delta i^* - \delta s^* - (\gamma + \delta r^*) \frac{\chi}{\sigma+\chi}$ $= \phi + (\gamma + \delta r^*) \left(1 - \frac{\chi}{\sigma+\chi} \right) = \phi + (\gamma + \delta r^*) \frac{\sigma}{\sigma+\chi}$ and $\rho + \mu + p + \delta + \gamma - \delta i^* - \beta^* s^* = \rho + \mu + p - \phi > 0$

so $M^* = \begin{bmatrix} \sigma + \beta^* i^* + \lambda^* \frac{\sigma}{\sigma+\chi} & -\beta^* i^* \\ \phi + (\gamma + \delta r^*) \frac{\sigma}{\sigma+\chi} & \rho + \mu + p - \phi \end{bmatrix}$. Equation (27) implies that the lower right

element of M^* is positive, so the sign pattern of M^* is $\begin{matrix} + & - \\ + & + \end{matrix}$, so $\det M^* > 0$. **Case II**:

Assume $i^* = 0$, so $M^* = \begin{bmatrix} \sigma + \lambda^* \frac{\sigma}{\sigma+\chi} & 0 \\ (\beta^* - \delta) s^* - (\gamma + \delta r^*) \frac{\chi}{\sigma+\chi} & \sigma + \delta + \gamma - \beta^* s^* \end{bmatrix}$. It suffices to

prove that $\sigma + \delta + \gamma - \beta^* s^* > 0$. Optimal policy implies that (β^*, λ^*) must lie in or on triangle ABC in Figure 2. Since $i^* = 0$ and $\beta = \beta^*$, it must be the case that $\lambda^* = \lambda_h (\beta^*)$ so Statement 2 of Proposition 5] implies $\beta^* s^* = \beta_c$. Therefore, $\sigma + \delta + \gamma - \beta^* s^* = \sigma + \delta + \gamma - \beta_c$ $= \sigma + \delta + \gamma - (\delta + \gamma + \phi) = \sigma - \phi = \rho + \mu + p - \phi > 0$, where the inequality follows from equation (27). ∎

Proof of Proposition 10. Suppose, contrary to what is to be proved, that the steady state under optimal policy is DFE so that $i^* = 0$. In Lemma 7, we show that $V_s - V_i > 0$ is finite. Therefore, if the steady state is DFE, $(V_s^* - V_i^*) i^* = 0$, so the first-order condition in equation (45) implies $y' (\beta_{OP}^*) = 0$, which implies $\beta_{OP}^* \geq \overline{\beta}$, so $\beta_r^* \geq \overline{\beta}$ and hence $y (\beta_r^*) = y (\overline{\beta})$. With $y (\beta_r^*) = y (\beta_{OP}^*) = y (\overline{\beta})$ and $(V_s^* - V_i^*) i^* = 0$, equation (47) becomes $V_r^* - V_s^* = \frac{\lambda^*}{\rho+\mu+\chi+\lambda^*} (1 - \zeta) \psi < \psi$, which contradicts the first-order condition in equation (46). [21] Therefore, the steady state under optimal policy cannot be DFE, which implies

---

[21]The first-order condition in equation (46), $(V_r - V_s) s \geq \psi (s + \zeta r)$ implies $V_r - V_s \geq \psi \left(1 + \zeta \frac{r}{s} \right) \geq \psi$

$\beta_{OP}^* > \beta_c$. To prove that $\beta_{OP}^* < \overline{\beta}$, suppose, contrary to what is to be proved, that $\beta_{OP}^* = \overline{\beta}$. Then Lemma 7 implies that $V_s^* - V_i^* > 0$. Since the steady state cannot be DFE, we have $i^* > 0$, which implies $s^* = R_0^{*-1} > 0$. Therefore, $(V_s^* - V_i^*) s^* i^* > 0 = (s^* + \zeta r^*) y'(\overline{\beta}) = (s^* + \zeta r^*) y'(\beta_{OP}^*)$, which violates the first-order condition in equation (45). ∎