Optimal mix of screening and contact tracing for endemic diseases[*]

Benjamin Armbruster[†]   Margaret L. Brandeau[†]

September 15, 2007

Abstract

Two common means of controlling infectious diseases are screening and contact tracing. Which should be used, and when? We consider the problem of determining the cheapest mix of screening and contact tracing necessary to achieve a desired endemic prevalence of a disease or to identify a specified number of cases. We perform a partial equilibrium analysis of small-scale interventions, assuming that prevalence is unaffected by the intervention; we develop a full equilibrium analysis where we compare the long-term cost of various combinations of screening and contact tracing needed to achieve a given equilibrium prevalence; and we solve the problem of minimizing the total costs of identifying and treating disease cases plus the cost of untreated disease cases. Our analysis provides several insights. First, contact tracing is only cost effective when prevalence is below a threshold value. This threshold depends on the relative cost per case found by screening versus contact tracing. Second, for a given contact tracing policy, the screening rate needed to achieve a given prevalence or identify a specified number of cases is a decreasing function of disease prevalence. As prevalence increases above the threshold (and contact tracing is discontinued), the screening rate jumps discontinuously to a higher level. Third, these qualitative results hold when we consider unchanged or changed prevalence, and short-term or long-term costs.

1 Introduction

Screening and contact tracing are two common means of controlling infectious diseases. Which is used depends on the disease and the context or region. For example, contact tracing (also known as partner notification) is routine for tuberculosis (TB) in the general population of low-prevalence countries. TB screening is often performed in higher prevalence contexts such as prisons and nursing homes.

Blanket screening for disease is uncommon. However, routine screening (either mandatory or recommended) for some diseases is performed in immigrants and high-risk groups. For example, immigrants are screened for TB and human immunodeficiency virus (HIV) in the US and Canada. The Centers for Disease Control and Prevention (CDC) recommend annual TB skin tests for prison employees and long-term inmates (CDC, 1996). TB screening is also commonly required for school employees and health care workers (CDC, 1993b). For pregnant women in the US, the CDC recommends screening for syphilis, Hepatitis B, chlamydia, gonorrhea, and Hepatitis C for those at high risk, and also recommends that pregnant women be offered an HIV test (CDC, 2002).

Contact tracing practice varies among diseases. In the US, health departments are legally required to perform contact tracing for all TB index cases (CDC, 2000). The CDC recommends routine contact tracing for HIV in the US (CDC, 2002), although it is not mandatory; often individuals testing positive for HIV do not want their partners informed for fear of discrimination (Pealer and Peterman, 2003). In Sweden, partner notification of HIV status is mandatory (Cowan et al., 1996). For STDs other than HIV, contact tracing has been practiced for decades (since the 19th century in Sweden and the UK and since the early 20th century in the US, for example) although it is not always routine (Cowan et al., 1996; Clarke, 1998). In one study cited in Clarke (1998), only 7–16% of chlamydia cases were referred for contact tracing.

A number of empirical studies of the cost effectiveness of contact tracing have been carried out. These include studies of TB (Dasgupta et al., 2000; Macintyre et al., 2000), HIV (Varghese et al., 1999; Cohen et al., 2004), and other STDs including chlamydia (Howell et al., 1997) and syphilis (Oxman and Doyle, 1996). For example, Oxman and Doyle (1996) performed an empirical study to compare the benefits of screening and contact tracing for early stage syphilis in Multnomah County, Oregon. The authors looked at the costs of screening and contact tracing, the number

of infected individuals found by contact tracing, and the number of infected individuals found by screening (including walk-ins, who see a practitioner for symptoms, and individuals found by random screening). They determined that the average cost of finding an infected individual was slightly lower among the screened individuals than among those found by contact tracing.

Some authors have developed detailed simulation models of the effectiveness of contact tracing. For example, Kretzschmar et al. (2004) and Porco et al. (2004) looked at the efficacy of contact tracing and ring vaccination to control a smallpox outbreak. Other authors (e.g., Müller et al. (2000) and Huerta and Tsimring (2002)) have developed theoretical models to evaluate the effectiveness of contact tracing. For example, Huerta and Tsimring (2002) used a stochastic network model to derive mean-field ordinary differential equations that describe epidemic quantities of interest, such as the number of individuals who are infected, susceptible, traced, and removed. None of these models examines the cost effectiveness of contact tracing. Closest to our approach is the multiple compartment gonorrhea transmission model of Hethcote and Yorke (1984). The authors compare the efficacy of different control strategies including contact tracing but do not model their costs.

The goal of our analysis is to determine the cheapest mix of screening and contact tracing necessary to achieve a desired endemic prevalence of a disease or to identify a specified number of cases. We present a mathematical framework for this problem. We perform a partial equilibrium analysis of small-scale interventions, assuming that prevalence is unaffected by the intervention, and then develop a full equilibrium analysis where we compare the long-term cost of various combinations of screening and contact tracing needed to achieve a given equilibrium prevalence. We also solve the problem of minimizing the total costs of screening, contact tracing, and treated and untreated disease cases.

2   Framework

In this section we present our mathematical framework. All notation is summarized in Table 1. We consider an infectious disease that is endemic in a population of size N consisting of S susceptible people and I infected people ($N = S + I$). We let $p := I/N$ denote the prevalence of the disease in this population. We assume that the modeled population includes only those people who can get the disease (for example, for STDs the population would not include young children).

We describe the dynamics of the disease using a standard random mixing model with susceptible and infected individuals, $S \rightarrow I$ (see, for example, Bailey (1975)). We assume that immunity to the disease does not exist (neither a vaccine nor previous infection provides immunity). This is reasonable for HIV, TB, and many STDs. We also assume a constant population size: the number of people entering the susceptible subgroup equals the number of people treated plus any who die in either the susceptible or infected subgroups. The assumption of a constant population size is realistic if relatively few people die of the disease.

Base Case   We first consider a base case in which there is no screening or contact tracing. New cases are identified via "walk-ins"; these are patients who see a health practitioner for symptoms. The change in prevalence at any point in time is described by the following set of differential equations:

$$
\begin{aligned}
\dot{S} &= - \left[\beta SI/N + \eta S \right] - \mu_1 S + \mu_2 I + \mathsf{walk\text{-}ins}(I) + \mu_1 N \\
\dot{I} &= + \left[\beta SI/N + \eta S \right] - \mu_1 I - \mu_2 I - \mathsf{walk\text{-}ins}(I).
\end{aligned}
\tag{1}
$$

The bracketed term represents the incidence of new cases: β is the sufficient contact rate and ηS the annual rate at which susceptible individuals in the population become infected via exogenous sources. This could be due, for example, to international travel of susceptible people who come back infected. Assuming net immigration is zero, the ηS term could also account for the arrival of infected immigrants. We assume that the per capita death rate in the population is μ_1 and the (untreated) recovery rate from the disease is μ_2. Note that the rate of entry into the population $\mu_1 N$ equals the total death rate $\mu_1(S + I)$. For diseases such HIV with no natural recovery, we can set $\mu_2 = 0$; for diseases such as TB where natural recovery can occur, we set $\mu_2 > 0$. In this system of equations, (infected) walk-ins arrive at an annual rate, $\mathsf{walk\text{-}ins}(I)$, that depends on the number of infected individuals. (We state all rates in annual terms, but any time unit could be used.)

We can rewrite the system of equations (1) as a single differential equation (since $N = S + I$), which we express in terms of prevalence, $p := I/N$:

$$
\dot{p} = \left[\beta p(1 - p) + \eta(1 - p) \right] - \mu p - \omega(p).
\tag{2}
$$

4

In equation (2), the term $\omega(p) := \mathsf{walk\text{-}ins}(I)/N$ denotes the per capita rate of walk-ins; and $\mu := \mu_1 + \mu_2$ so that μp is the reduction in prevalence due to deaths or recovery of infected individuals. For the remainder of our analysis, the fundamental quantity we examine is the per capita prevalence p. (Absolute numbers of infected and susceptible individuals can be readily calculated from this quantity.)

We model the annual rate of walk-ins as a fraction, γ, of the incidence of new cases:

$$\omega(p) := \gamma(1 - p)[\beta p + \eta]. \tag{3}$$

We assume that $\gamma < 1$; not all infected individuals seek treatment. Modeling the effect of walk-ins in this way is an approximation since in reality there is some delay between the time an individual becomes infected, develops symptoms, and is then treated or removed from the population. The yearly cost per capita to identify and treat new cases is

$$\omega(p)C_t \tag{4}$$

where C_t is the cost of confirmatory tests, counseling, and treatment for a single case of disease.

Screening and Contact Tracing We consider two interventions to identify (and subsequently treat) new cases: screening and contact tracing. We assume that screening can perfectly discriminate between infected and uninfected individuals; this is a reasonable assumption if screening comprises a battery of tests that taken together have near perfect sensitivity and specificity. Below we extend the model to allow for imperfect screening. We assume that screening occurs at an annual rate of λN people per year and finds infected individuals at a per capita annual rate of λp. We assume that the baseline cost per person screened is C_S. This does not include the cost of confirmatory tests, treatment, or counseling for those who test positive.

Contact tracing, if implemented, is performed for all walk-ins and infected screened individuals; these are the index cases. We model the presence or absence of contact tracing as a binary variable, δ, that takes value 1 if contact tracing is present and 0 otherwise. We assume that the mean number of additional infected individuals per index case found via contact tracing is K_T, and that the total cost of finding, testing, and treating these K_T individuals is C_T. We assume that the cost per

infected person found and treated via contact tracing is greater than the usual treatment cost: thus, $C_T/K_T > C_t$.

With (perfect) screening and contact tracing, the cost function becomes

$$C(\lambda, \delta; p) := \omega(p)C_t + \lambda(C_S + C_t p) + \delta C_T(\lambda p + \omega(p)), \tag{5}$$

and the rate of change in prevalence is given by

$$\dot{p} = f(p, \lambda, \delta) := (1 - p)[\beta p + \eta] - \mu p - (1 + \delta K_T)(\lambda p + \omega(p)). \tag{6}$$

We model imperfect screening by assuming that an imperfect initial test is followed by a perfect (more expensive) test for those who initially test positive. Then screening policy λ finds $\lambda s_{\mathrm{tp}} p$ infected individuals, where s_{tp} is the sensitivity (true positive rate) of the test; and yields positive initial tests for $\lambda s_{\mathrm{fp}}(1 - p)$ uninfected individuals, where s_{fp} is the false positive rate of the test. The cost of confirmatory tests for individuals with true positive tests is already included in the treatment cost C_t. We assume that the cost of followup screening tests for an individual with a false positive test is C_{fS}. With these assumptions, equations (5) and (6) become

$$C(\lambda, \delta; p) := \omega(p)C_t + \lambda(C_S + C_t s_{\mathrm{tp}} p + C_{\mathrm{fS}} s_{\mathrm{fp}}(1 - p)) + \delta C_T(\lambda s_{\mathrm{tp}} p + \omega(p)), \tag{7}$$

$$\dot{p} = f(p, \lambda, \delta) := (1 - p)[\beta p + \eta] - \mu p - (1 + \delta K_T)(\lambda s_{\mathrm{tp}} p + \omega(p)). \tag{8}$$

Note that for a given endemic prevalence p there are only two steady-state policies, that is, policies that remain the same for all time: $(\lambda = \lambda_0, \delta = 0)$ and $(\lambda = \lambda_1, \delta = 1)$. Thus, to determine a minimum cost steady-state policy, we determine the optimal screening level with and without contact tracing and then compare the costs of the two policies.

We now consider three different formulations of the problem: one in which screening and contact tracing do not change the disease prevalence (Section 3); one in which screening and contact tracing do change the disease prevalence, and the goal is to minimize the long-term annual cost of maintaining a specified equilibrium prevalence (Section 4); and one in which screening and contact tracing change the disease prevalence, and the goal is to determine the level of control (and thus

Decision Variables

λ	per capita screening rate
δ	1 if contact tracing is performed, 0 otherwise

Optimal Policy

$(\lambda_U^*, \delta_U^*)$	optimal policy for the case of unchanged prevalence
$(\lambda_L^*, \delta_L^*)$	optimal policy for the case of changed prevalence, and minimization of long-term cost
$(\lambda_T^*, \delta_T^*)$	optimal policy for the case of changed prevalence, and minimization of total cost

Costs

$C(\lambda, \delta; p)$	total annual per capita cost of screening policy λ and contact tracing policy δ
C_S	initial cost per person screened
C_{fS}	cost of followup screening for a false positive
C_t	cost of confirmatory tests, counseling, and treating a single case of disease identified via screening or walk-in
C_T	contact tracing cost per index case

Other Notation

I	number of infected individuals in the population
S	number of susceptible individuals in the population
N	population size, $N = I + S$
p	prevalence, $p = I/N$
$f(p, \lambda, \delta)$	rate of change in prevalence
β	sufficient contact rate
γ	fraction of incident cases who are walk-ins (who see a health practitioner for symptoms)
η	per capita exogenous rate of new infections (from sources external to the population)
K_T	number of secondary cases per index case found by contact tracing
μ_1	per capita natural death rate in the population
μ_2	per capita untreated recovery rate from the disease
μ	per capita exit rate from infected population due to death or recovery, $\mu = \mu_1 + \mu_2$
walk-ins(I)	annual rate of walk-ins
$\omega(p)$	per capita walk-in rate, $\omega(p) = \text{walk-ins}(pN)/N$
n	number of new cases to be found per year
α	cost per year of one infection
r	discount rate used to discount future costs
s_{tp}	true positive rate of initial screening
s_{fp}	false positive rate of initial screening

the associated equilibrium disease prevalence) that minimizes the total costs of screening, contact tracing, and disease (Section 5). We assume in the first formulation (Section 3) that the population is large enough and prevalence is low enough such that the disease control efforts do not change the disease prevalence; this might be the case, for example, if the disease is endemic (e.g., chlamydia) and the budget for disease control is low. The second formulation (Section 4) could reflect a situation in which the budget is high enough to reduce disease prevalence and the public health department takes a long-term perspective, aiming to minimize only the long-term costs of control. The third formulation (Section 5) could reflect the situation in which the budget is high enough to reduce disease prevalence and the health department takes a shorter-term perspective.

3   Unchanged prevalence

Suppose that screening and contact tracing have a small enough effect that the disease prevalence p remains (essentially) the same. Then the problem of determining the optimal mix of screening versus contact tracing can be viewed as one of finding and treating a specified number of cases per year at minimum cost. Building on equations (7) and (8), let $C_U^*(n; p)$ be the minimum cost of finding and treating at least n cases per year when the prevalence is p. We assume that $n > N\omega(p)$: we want to find more infected individuals than just the walk-ins. In one year at prevalence p we find $N(1 + \delta K_T)(\lambda s_{\mathrm{tp}} p + \omega(p))$ cases with screening rate λ and contact tracing policy δ. The minimum cost is then given by

$$C_U^*(n; p) := \min_{\delta, \lambda} \ C(\lambda, \delta; p)$$
$$\text{s.t. } N(1 + \delta K_T)(\lambda s_{\mathrm{tp}} p + \omega(p)) \geq n \tag{9}$$
$$\lambda \geq 0, \quad \delta \in \{0, 1\}.$$

Let $(\lambda_U^*(n; p),\ \delta_U^*(n; p))$ denote the optimal policy. Since the constraint is linear in λ and the objective function is increasing in λ, we obtain

$$\lambda_U^*(n; p) = \frac{1}{N p s_{\mathrm{tp}}} \left[\frac{n}{1 + \delta_U^*(n; p) K_T} - N\omega(p) \right]^+ \tag{10}$$

where $[x]^+ := \max(x, 0)$ is the positive part operator. When the bracketed term in equation (10) is negative, then we find more than n cases just by treating the walk-ins (and their contacts if $\delta = 1$). Because we wish to focus on the trade-off between screening and contact tracing we assume that we want to find more cases than just the walk-ins and their contacts; thus, we henceforth assume that $n \geq N\omega(p)(1 + K_T)$.

Substituting (10) into (9), we obtain

$$C_U^*(n; p) = \frac{1}{Nps_{\text{tp}}} \left[n \frac{C_S + C_t s_{\text{tp}} p + C_{\text{fS}} s_{\text{fp}}(1-p) + ps_{\text{tp}}\delta_U^*(n; p)C_T}{1 + \delta_U^*(n; p)K_T} - N\omega(p)(C_S + C_{\text{fS}} s_{\text{fp}}(1-p)) \right].$$

(11)

Comparing $C_U^*(n; p)$ for $\delta = 0$ and $\delta = 1$ we find that the optimal policy for contact tracing is

$$\delta_U^*(n; p) = \begin{cases} 1 & \text{if } p < \frac{C_S + C_{\text{fS}} s_{\text{fp}}}{s_{\text{tp}}C_T/K_T + C_{\text{fS}} s_{\text{fp}} - C_t s_{\text{tp}}}, \\ 0 & \text{otherwise.} \end{cases}$$

(12)

These expressions simplify in the case of perfect screening (the terms with s_{fp} drop out). For example, the condition in (12) becomes $p < \frac{C_S}{C_T/K_T - C_t}$.

From (12), we see that contact tracing is optimal when the prevalence is below a threshold, regardless of the number of cases we want to find. This is not unexpected: threshold results have been found when comparing various combinations of screening, testing, and vaccination strategies for other diseases, e.g., Hutton et al. (2007); Rubió (1997); Owens et al. (1996).

The contact tracing threshold is a positive quantity because of our assumption that $C_T/K_T > C_t$. We denote this threshold prevalence by P'. In the case of perfect screening, $C_S/P' = C_T/K_T - C_t$, meaning intuitively that at prevalence P' the marginal cost of finding an additional case via screening, C_S/P', equals the average cost of finding a new case via contact tracing, $C_T/K_T - C_t$ (we subtract C_t because C_T includes the cost of treating infected contacts while we are interested only in the cost of *finding* them).

From (10), we see that the optimal screening rate increases with the number of cases to be found (n). The optimal screening rate decreases if contact tracing (δ) is implemented. Additionally, the optimal screening rate decreases as the walk-in rate ($\omega(p)$) increases.

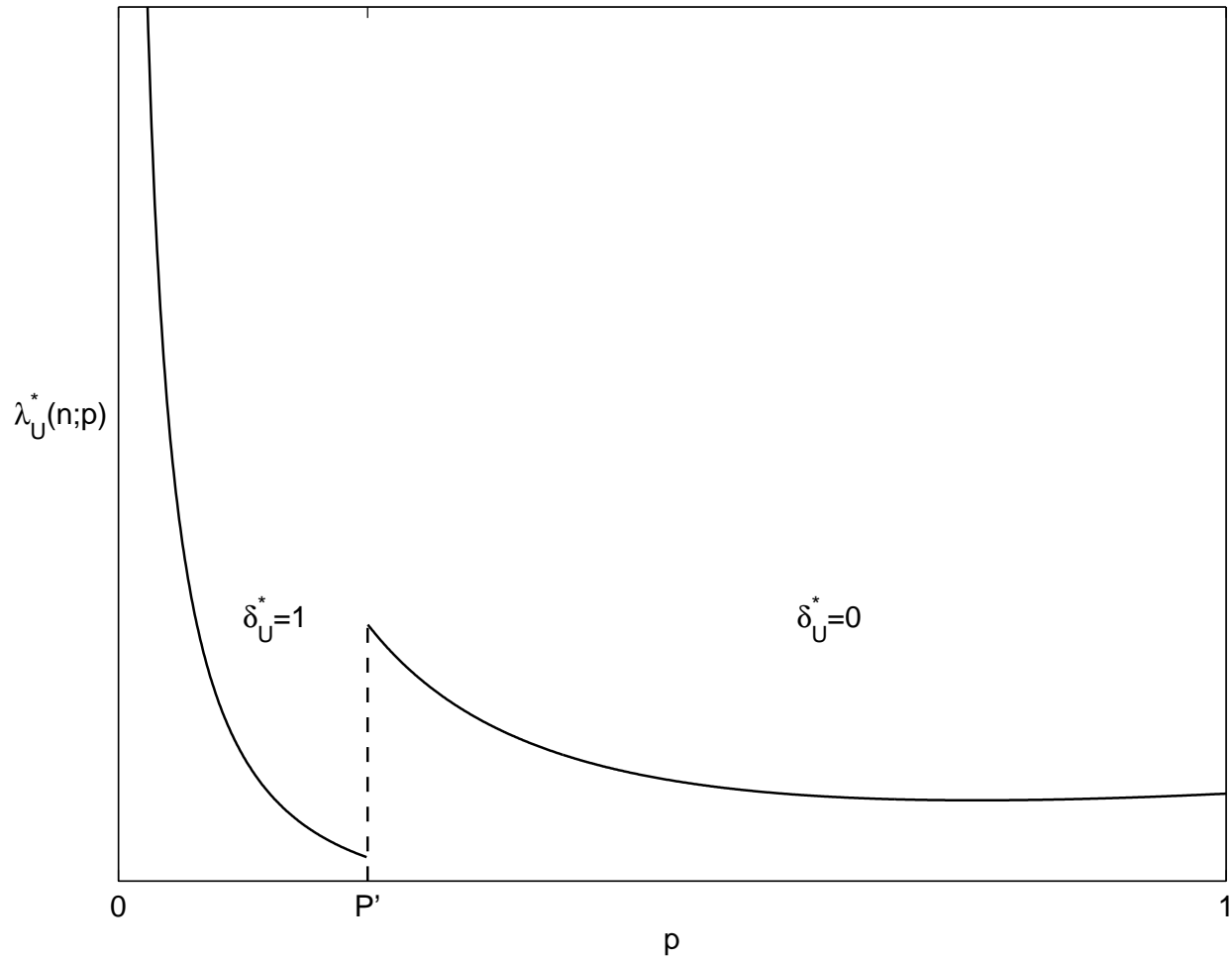Figure 1 shows schematically how the optimal amount of screening needed to identify n cases

9

Figure 1: Optimal levels of screening and contact tracing to find n cases as a function of existing endemic disease prevalence p, assuming that prevalence is unchanged by screening and contact tracing

$C_U^*(n;p)$

$\delta_U^*=1$
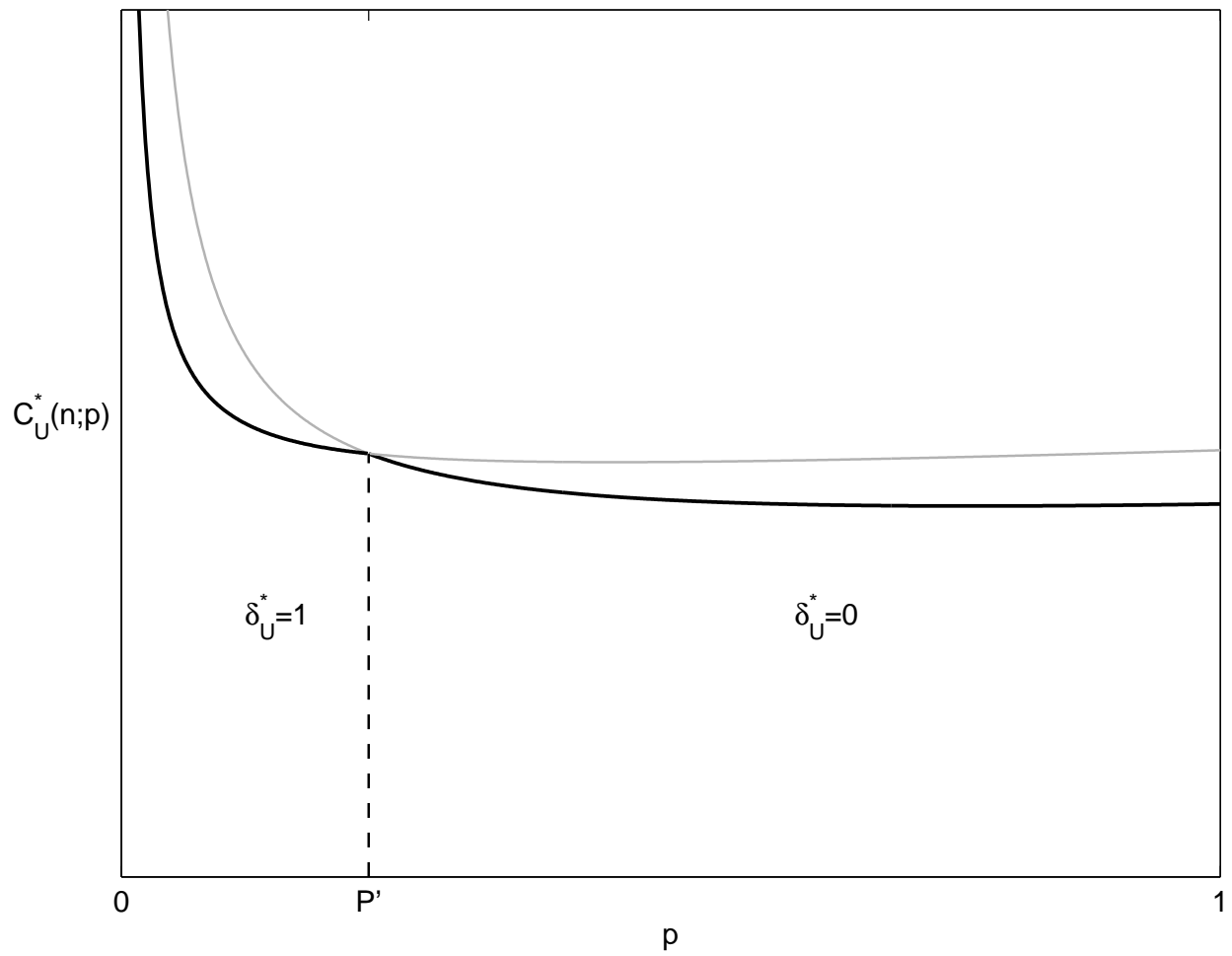
$\delta_U^*=0$

0

P'

1

p

Figure 2: Minimum cost to find n cases as a function of existing endemic disease prevalence p, assuming that prevalence is unchanged by screening and contact tracing

varies with the endemic disease prevalence. This figure, and all subsequent figures in this paper, use the following parameter values: $s_{\text{tp}} = 1$, $K_T = 1.5$, $\gamma = 0.1$, $\beta = 0.15$, $\eta = 0.01$, $\mu = 0.05$, $s_{\text{fp}} = 0$, $C_{\text{fS}} = 50$, $C_t = 200$, $C_S = 15$, $C_T = 400$, $n = 10$, and $N = 1000$. We note that with other parameter values, all figures remain qualitatively similar to those shown. When the prevalence is below P', contact tracing is used, and the optimal screening level decreases as prevalence increases. At the threshold P', contact tracing is discontinued, and the optimal level of screening jumps discontinuously to a higher level. For prevalence above P', the optimal screening rate decreases in proportion to $1/p$. Eventually, as p gets large enough, the optimal screening rate begins to increase slightly. This is because the number of walk-ins, $N\omega(p)$, decreases due to saturation (note the $p(1 - p)$ term in $\omega(p)$).

Figure 2 shows the cost of the optimal screening policy (11), with and without contact tracing. These costs are to find a fixed number, n, of infected individuals and therefore decrease with p (except for very high prevalences). The cost decreases with higher prevalence because there are more walk-ins (except for very high prevalences where the disease saturates the population) and more infected individuals, and therefore less screening is needed to find the n individuals. The cost of the optimal screening and contact tracing policy is the lower envelope of the two curves. We observe that, as prevalence increases, screening becomes more effective, and the marginal cost per infected found, $\partial C_U^* / \partial n$, decreases:

$$\frac{\partial C_U^*(n; p)}{\partial p \partial n} = \frac{-1}{p^2} \frac{C_S + C_{\text{fS}} s_{\text{fp}}}{1 + \delta_U^*(n; p) K_T} \frac{1}{N s_{\text{tp}}}. \tag{13}$$

4 Minimizing Long-term Costs

Under the assumptions of our model (8), if the same policy (λ, δ) is used for an extended period, then prevalence $p(t)$ will converge to some steady-state value, $P \leq P_0$. Here P_0 is the saturation prevalence with neither screening nor contact tracing. Long-term disease control costs — that is, ignoring transition costs — are determined by the steady-state prevalence, P, that is achieved. Hence it is useful to determine for each steady-state prevalence $P \leq P_0$ the cheapest policy that

maintains that prevalence, and the cost of that policy. We write this problem as:

$$C_L^*(P) := \min_{\delta,\lambda} C(\lambda, \delta; p)$$

$$s.t. \ \dot{p}(t) = 0, \quad p(t) = P \leq P_0 \ \forall t \tag{14}$$

$$\lambda \geq 0, \quad \delta \in \{0, 1\} \ \forall t.$$

Substituting (8) into (14) yields

$$C_L^*(P) := \min_{\delta,\lambda} C(\lambda, \delta; P)$$

$$s.t. \ 0 = (1 - P)[\beta P + \eta] - \mu P - (1 + \delta K_T)(\lambda s_{\text{tp}} P + \omega(P)) \tag{15}$$

$$\lambda \geq 0, \quad \delta \in \{0, 1\}.$$

The value of P_0, obtained by substituting $\delta = 0$ and $\lambda = 0$ into the constraint of (15) and choosing the larger solution, is

$$P_0 = \frac{1}{2} - \frac{\eta}{2\beta} - \frac{\mu}{2\beta(1 - \gamma)} + \sqrt{\left(\frac{1}{2} - \frac{\eta}{2\beta} - \frac{\mu}{2\beta(1 - \gamma)}\right)^2 + \frac{\eta}{\beta}} \tag{16}$$

which equals $1 - \frac{\mu}{\beta(1-\gamma)}$ if $\eta = 0$.

To optimize (14) we compare the two steady-state policies, $(\lambda = \lambda_0, \delta = 0)$ and $(\lambda = \lambda_1, \delta = 1)$, for an endemic prevalence of P. Using analysis similar to that in Section 3, we solve for λ in the constraint of (15):

$$\lambda_L^*(P) = \frac{1}{P s_{\text{tp}}} \left[\frac{\beta P(1 - P) + \eta(1 - P) - \mu P}{1 + \delta_L^*(P) K_T} - \omega(P) \right], \tag{17}$$

which can be expressed as

$$\lambda_L^*(P) = \frac{1}{P s_{\text{tp}}} \frac{(1 - \gamma(1 + \delta_L^*(P) K_T))[\beta P(1 - P) + \eta(1 - P)] - \mu P}{1 + \delta_L^*(P) K_T}. \tag{18}$$

The numerator in (18) is the number of new cases each year that we need to find to maintain disease prevalence P (incident cases minus walk-ins and deaths). We divide by $1 + \delta K_T$ to determine the equivalent number of index cases, and by $P s_{\text{tp}}$ to determine the number of people we need to screen.

As expected, the numerator (and hence λ_L^*) is nonnegative for $P \leq P_0$ (where P_0 is given by (16)). Substituting (17) into (7) we obtain

$$C_L^*(P) = \omega(P)(C_t + \delta_L^*(P)C_T)$$
$$+ \left[\frac{\beta P(1 - P) + \eta(1 - P) - \mu P}{1 + \delta_L^*(P)K_T} - \omega(P) \right] \frac{C_S + C_t s_{tp} P + C_{fS} s_{fp}(1 - P)}{P s_{tp}}. \quad (19)$$

Figure 3 shows the optimal screening rate as a function of the desired endemic disease prevalence, P, and Figure 4 shows the associated cost. The shape of the cost function in Figure 4 can be explained as follows. The cost function, (7) (upon which (19) is based), contains two counteracting terms: the rate of walk-ins, $\omega(p)$, and the screening costs λp. The walk-in rate is quadratic in p, increasing for small prevalence and decreasing near saturation. So for a constant screening level λ, costs tend to rise at small prevalences and then decrease later. However, in the long-term cost setting, the optimal screening level, $\lambda_L^*(P)$, varies with prevalence, (18). If there were no infections from outside the population ($\eta = 0$) then the optimal screening level $\lambda_L^*(P)$ would be linear in prevalence and the cost would remain quadratic in P (first increasing and then decreasing). For $\eta > 0$, however, costs change in proportion to $1/P$ for very low prevalence because the $\eta(1 - P)/(P s_{tp})$ term dominates. Thus, costs decrease sharply for very small P. Beyond some value of P (the local minimum between 0 and P' in Figure 4) the impact of outside infections, η, is negligible and the cost behaves quadratically in P (first increasing and then decreasing).

Comparing C_L^* for $\delta = 0$ and $\delta = 1$ we find that the optimal policy for contact tracing is

$$\delta_L^*(P) = \begin{cases} 1 & \text{if } P < \frac{C_S + C_{fS} s_{fp}}{s_{tp} C_T / K_T + C_{fS} s_{fp} - C_t s_{tp}}, \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

just as in Section 3 in the unchanged prevalence setting. (There is a second unphysical $P > P_0$ where the costs, C_L^*, in (19) are the same with and without contact tracing.)
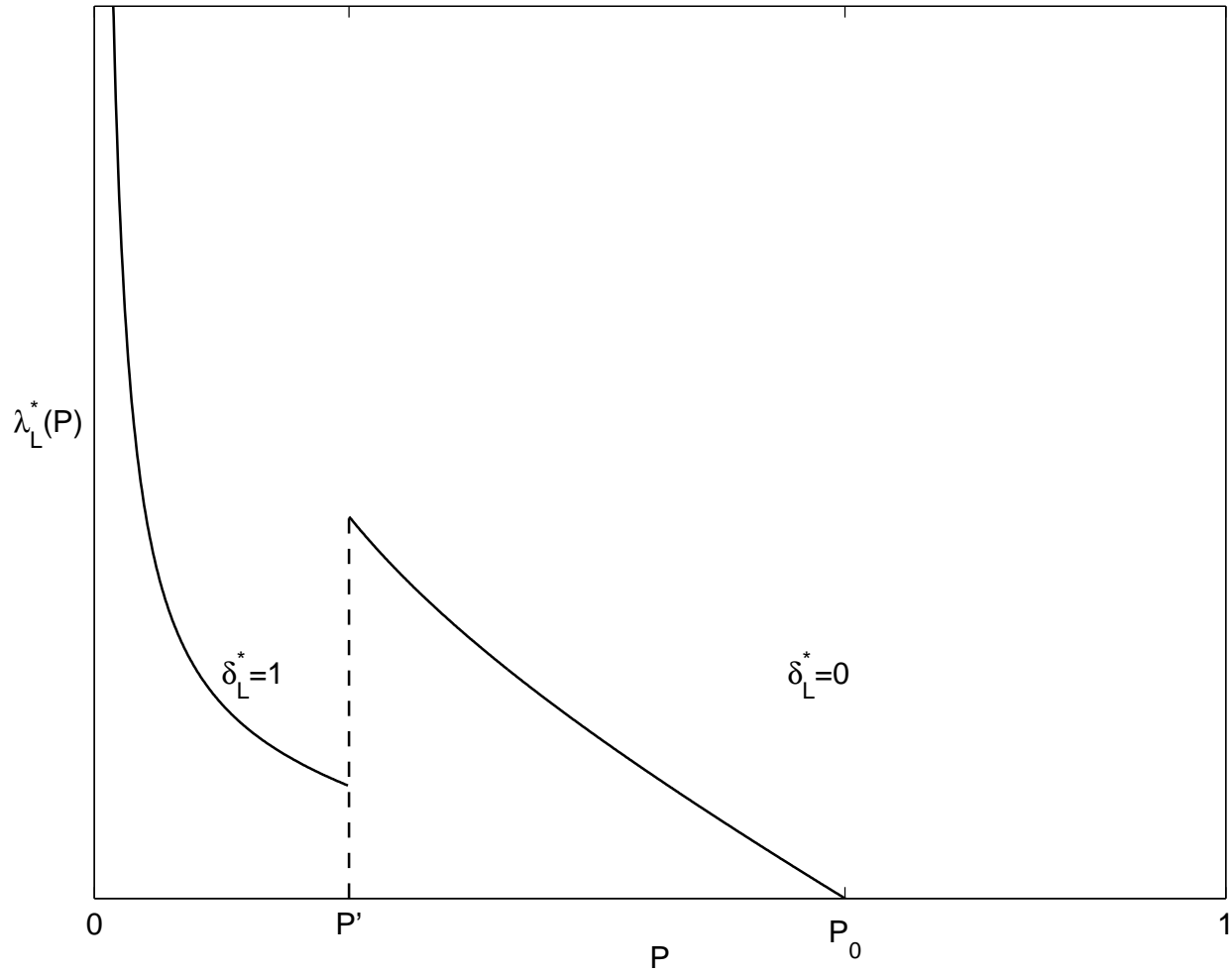
Figure 3: Optimal screening level λ_L^* and contact tracing policy δ_L^* as a function of the desired endemic disease prevalence P, when the goal is to minimize long-term cost
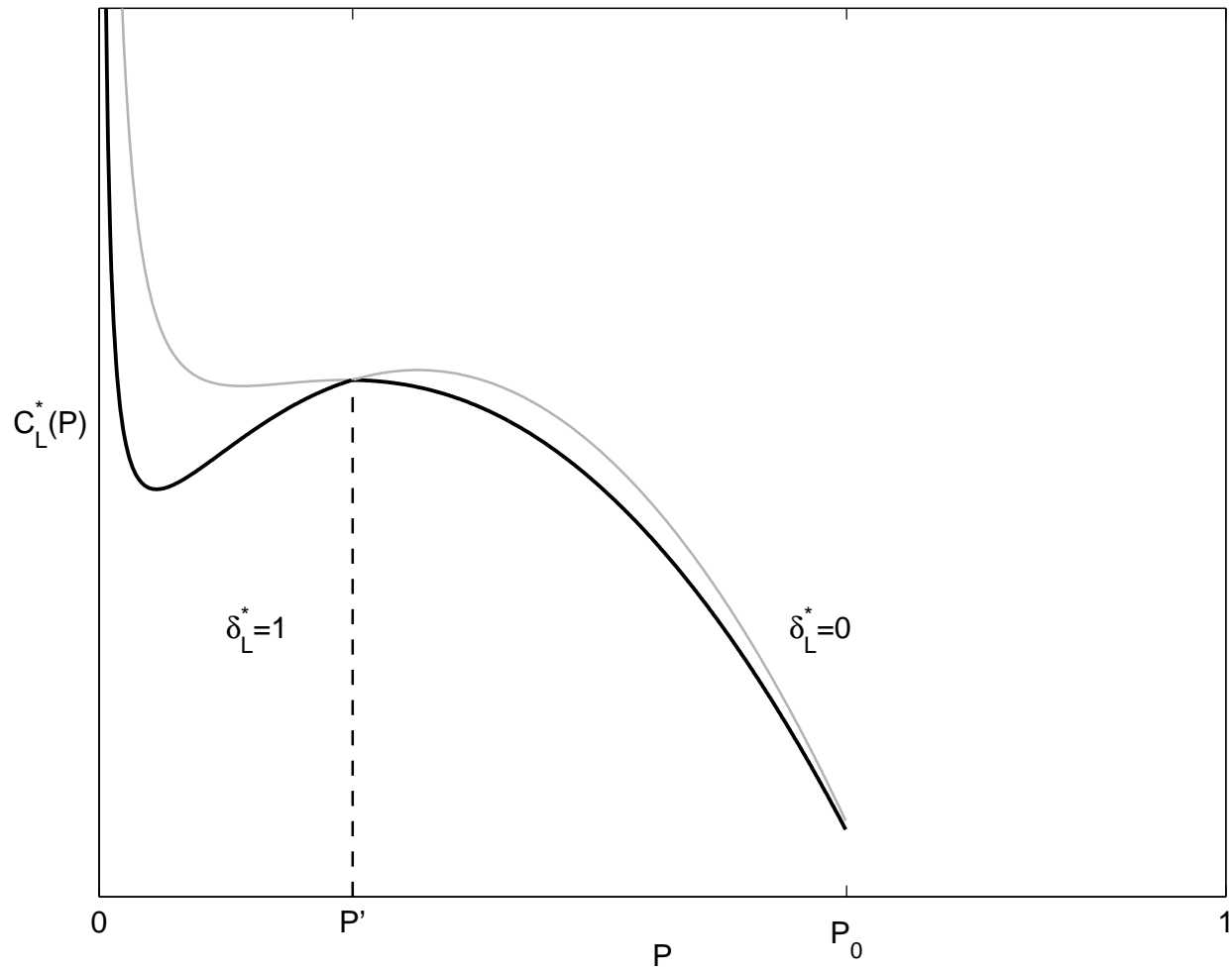
Figure 4: Minimum long-term cost as a function of the desired endemic disease prevalence P

5 Minimizing Total Costs

In the previous sections, we assumed that the disease prevalence was specified exogenously: either it was unchanged by screening and contact tracing (Section 3) or the long-term prevalence was specified and the goal was to determine the control policy to achieve that prevalence at minimum cost (Section 4). In this section we consider the problem of determining the control policy that minimizes the total cost of screening, contact tracing, and (treated and untreated) disease. The solution to this problem specifies not only the optimal screening and contact tracing policy, but also the cost-minimizing level of disease prevalence.

The cost function (7) includes the costs of screening, contact tracing, and treating identified cases of disease. To assign a cost to untreated disease cases (i.e., disease prevalence), we introduce a parameter α, which we define as the cost per year of one infection. To estimate α, we can apply concepts from cost-benefit analysis (Gold et al., 1996) and multiply the monetary value of one year of healthy life by the quality decrement associated with infection. For example, if a year of healthy life is worth \$50,000 (Owens, 1998) and the quality adjustment for infection is 0.9 (thus, the quality decrement is 0.1 healthy years lost per year of life with infection), then we would estimate α as \$5,000.

We state the problem as follows: given that the current prevalence is p_0, what is the best infinite-horizon dynamic policy $(\lambda(t), \delta(t))$ that minimizes the total cost of screening, contact tracing, and disease, while never allowing prevalence to increase? We assume a discount rate r, and use equation (8) as the state equation. We also assume a bound on the resources available at any point in time. This bound could reflect budget constraints or physical limits on the number of people who can be screened annually. Specifically, we assume a bound of the form $\lambda(t) \leq \bar{\lambda}$ where $\bar{\lambda}$ is some constant. We note that it should not be hard to extend the analysis to a more general bound $R(p(t), \lambda(t), \delta(t)) \leq B$ where $R(\cdot)$ is an increasing function of λ. We write the optimal control

problem as:

$$C_T^*(p_0, \alpha) := \min_{p(t),\delta(t),\lambda(t)} \int_0^\infty e^{-rt}[C(\lambda(t), \delta(t); p(t)) + \alpha p(t)]\, dt$$

$$\text{s.t.} \quad p(0) = p_0, \ \dot{p}(t) = f(p(t), \lambda(t), \delta(t)) \ \forall t \tag{21}$$

$$\dot{p}(t) \leq 0, \ \lambda(t) \leq \bar{\lambda} \ \forall t$$

$$0 \leq \lambda(t), \ \delta(t) \in \{0, 1\} \ \forall t.$$

This control problem seeks to minimize the discounted per capita cost of finding and treating disease, $\int_0^\infty e^{-rt}C(\cdot)$, plus the discounted per capita cost of (unidentified) untreated disease, $\int_0^\infty e^{-rt}\alpha p(t)$. We observe that as the level of disease control increases, the cost $\int_0^\infty e^{-rt}C(\cdot)$ increases and the cost $\int_0^\infty e^{-rt}\alpha p(t)$ decreases. Thus, the optimal level of disease control is such that the marginal cost of additional case detection equals the marginal cost of case non-detection.

Our optimal control problem is autonomous and the optimal trajectory has no time dependence: if an optimal trajectory associated with initial prevalence x has prevalence and policy $(p^*(\cdot; x), \lambda_T^*(\cdot; x), \delta_T^*(\cdot; x)) = \vec{g}(\cdot)$, then then for any t the function $s \mapsto g(s + t)$ gives an optimal prevalence and policy $(p^*(\cdot; g(t)), \lambda_T^*(\cdot; g(t)), \delta_T^*(\cdot; g(t)))$ associated with initial prevalence $g(t)$. This can be seen by decomposing our optimal control problem (21) for $t > 0$,

$$C_T^*(p_0, \alpha) := \min_{\substack{p(s),\delta(s),\lambda(s) \\ s \leq t}} C_T^*(p(t), \alpha) + \int_0^t e^{-rs}[C(\lambda(s), \delta(s); p(s)) + \alpha p(s)]\, ds$$

$$\text{s.t.} \quad p(0) = p_0, \ \dot{p}(s) = f(p(s), \lambda(s), \delta(s)) \ \forall s \leq t \tag{22}$$

$$\dot{p}(s) \leq 0, \ \lambda(s) \leq \bar{\lambda} \ \forall s \leq t$$

$$0 \leq \lambda(s), \ \delta(s) \in \{0, 1\} \ \forall s \leq t.$$

This means that the optimal policy $(\lambda_T^*(p; \alpha), \delta_T^*(p; \alpha))$ depends only on the current prevalence.

By the Pontryagin maximum principle, the optimal screening policy $\lambda_T^*(t)$ minimizes the Hamiltonian (with Lagrange multiplier ϕ) at every point along the optimal trajectory $p^*(t)$:

$$\text{for all } t, \quad \lambda_T^*(t) \in \arg\min_\lambda H(p^*(t), \phi^*(t), \lambda, t),$$

$$\text{where} \quad H(p, \phi, \lambda, t) := e^{-rt}[C(\lambda, \delta_T^*(p; \alpha); p) + \alpha p] + \phi f(p, \lambda, \delta_T^*(p; \alpha)). \tag{23}$$

Since the Hamiltonian $H(\cdot)$ at any time t is linear in λ (see (7) and (8)), it follows that the optimal control will be bang-bang: $\lambda_T^*(t)$ is either at its minimal level, $\underline{\lambda}(p; \alpha)$ or at its maximal level, $\bar{\lambda}$. Due to the constraint $\dot{p} = f(p, \lambda, \delta) \leq 0$ and the fact that $f(\cdot)$ is decreasing in λ, the minimal level is

$$\underline{\lambda}(p; \alpha) = \frac{1}{p s_{\text{tp}}} \left[\frac{\beta p(1-p) + \eta(1-p) - \mu p}{1 + \delta_T^*(p; \alpha) K_T} - \omega(p) \right]. \tag{24}$$

Let $\mathcal{F}(\alpha) := \{p : \underline{\lambda}(p; \alpha) \leq \bar{\lambda}\}$ be the feasible set of prevalences: those prevalences for which there exists some screening level satisfying both constraints. This set consists of one or two intervals: note the shape of Figure 3 and the similarity of $\underline{\lambda}(\cdot; \alpha)$ to $\lambda_L^*(\cdot)$. We assume the initial prevalence is feasible, $p_0 \in \mathcal{F}(\alpha)$, as otherwise (21) has no solution.

Since our problem is autonomous, the optimal contact tracing policy is the indicator function, $\delta_T^*(p; \alpha) = \mathbf{1}_{A(\alpha)}(p)$, of some set $A(\alpha)$. Using a calculation similar to those in the previous sections, for any screening level we can find a smaller screening level that decreases prevalence at the same rate when combined with contact tracing: we can find $\bar{\lambda}_1(p) \leq \bar{\lambda}$ and $\underline{\lambda}_1(p; \alpha) \leq \underline{\lambda}(p; \alpha)$ such that $f(p, \bar{\lambda}_1(p), 1) = f(p, \bar{\lambda}, 0)$ and $f(p, \underline{\lambda}_1(p; \alpha), 1) = f(p, \underline{\lambda}(p; \alpha), 0)$. In either case, the screening level with contact tracing is cheaper for prevalences $p \in [0, P')$: $C(p, \bar{\lambda}_1(p), 1) < C(p, \bar{\lambda}, 0)$ and $C(p, \underline{\lambda}_1(p; \alpha), 1) < C(p, \underline{\lambda}(p; \alpha), 0)$. Hence it is optimal to contact trace when prevalence is below a threshold, $[0, P') \subseteq A(\alpha)$.

The above considerations show that it is optimal to screen at the maximal rate, $\bar{\lambda}$, until prevalence reaches some level p_1 at time t_1 and then to remain at prevalence p_1. We can separate the costs of such a trajectory into the long-term costs of remaining at prevalence p_1,

$$\frac{C_L^*(p_1) + \alpha p_1}{r} = \int_0^\infty e^{-rt} [C_L(p_1) + \alpha p_1] \, \mathrm{d}t, \tag{25}$$

and the transition costs associated with changing the prevalence from p_0 to p_1,

$$J(p_0, p_1, \alpha) = \int_0^{t_1} e^{-rt} [C(\bar{\lambda}, \mathbf{1}_{A(\alpha)}(p); p(t)) - C_L(p_1) + \alpha(p(t) - p_1)] \, \mathrm{d}t \tag{26}$$

$$\dot{p}(t) = f(p(t), \bar{\lambda}, \mathbf{1}_{A(\alpha)}(p)), \quad p(0) = p_0, \quad p(t_1) = p_1.$$

The total cost is their sum, $J(p_0, p_1, \alpha) + (C_L^*(p_1) + \alpha p_1)/r$, and given $A(\alpha)$, our problem is then

to choose the p_1 that minimizes the total cost:

$$C_T^*(p_0, \alpha) = \min_{p_1 \leq p_0} J(p_0, p_1, \alpha) + (C_L^*(p_1) + \alpha p_1)/r. \tag{27}$$

We write $p_1^*(p_0; \alpha)$ for the value of p_1 that achieves the optimum. Note that as we discount less, $r \to 0$, only the long-term costs matter; that is $p_1^*(p_0; \alpha) \in \arg\min_{p \leq p_0} C_L(p) + \alpha p$. Additionally, as we constrain λ more (i.e., decrease $\bar{\lambda}$), then p_1^* cannot decrease (a proof is in the appendix).

When the screening rate is unbounded ($\bar{\lambda} = \infty$), our analytic tools go further because the optimal policy decreases prevalence in a single instantaneous jump at the beginning. In this case, all prevalences are feasible, $\mathcal{F}(\alpha) = [0, \infty)$; the optimal contact tracing policy is the same as in the previous sections, $A(\alpha) = [0, P')$; and $J(\cdot)$ is given by

$$J(p_0, p_1, \alpha) = \frac{C_t s_{\text{tp}} - C_{\text{fS}} s_{\text{fp}} + \delta_T^*(p_0) s_{\text{tp}} C_T}{(1 + \delta_T^*(p_0) K_T) s_{\text{tp}}}(p_0 - p_1) + \frac{C_S + C_{\text{fS}} s_{\text{fp}}}{(1 + \delta_T^*(p_0) K_T) s_{\text{tp}}} \log \frac{p_0}{p_1}, \tag{28}$$

when p_0 and p_1 are on the same side of P', the threshold prevalence for contact tracing. Equation (28) is derived in the appendix. If p_0 and p_1 are on opposite sides of P', $p_1 < P' < p_0$, then we calculate the transition cost as the cost of reducing prevalence (instantaneously) from p_0 to P' (with no contact tracing) plus the cost of reducing prevalence (instantaneously) from P' to p_1 (with contact tracing): $J(p_0, P', \alpha) + J(P', p_1, \alpha)$ where $\delta = 0$ in the $J(p_0, P', \alpha)$ term and $\delta = 1$ in the $J(P', p_1, \alpha)$ term.

For the example problem worked out in the previous two sections, with the additional parameters $r = 0.05$ and $\bar{\lambda} = \infty$, Figure 5 shows the value of p_1^* as a function of α, assuming that the constraint $p_1 \leq p_0$ is not binding. The optimal solution to this problem with the constraint $p_1 \leq p_0$ (thus the solution to the general control problem (21)) is the smaller of p_0 and the value shown in the figure. Figure 5 illustrates the relationship between p_1^* and α: as the cost of untreated disease increases, the optimal long-term prevalence decreases – which means that the optimal level of disease control (screening and contact tracing) increases.
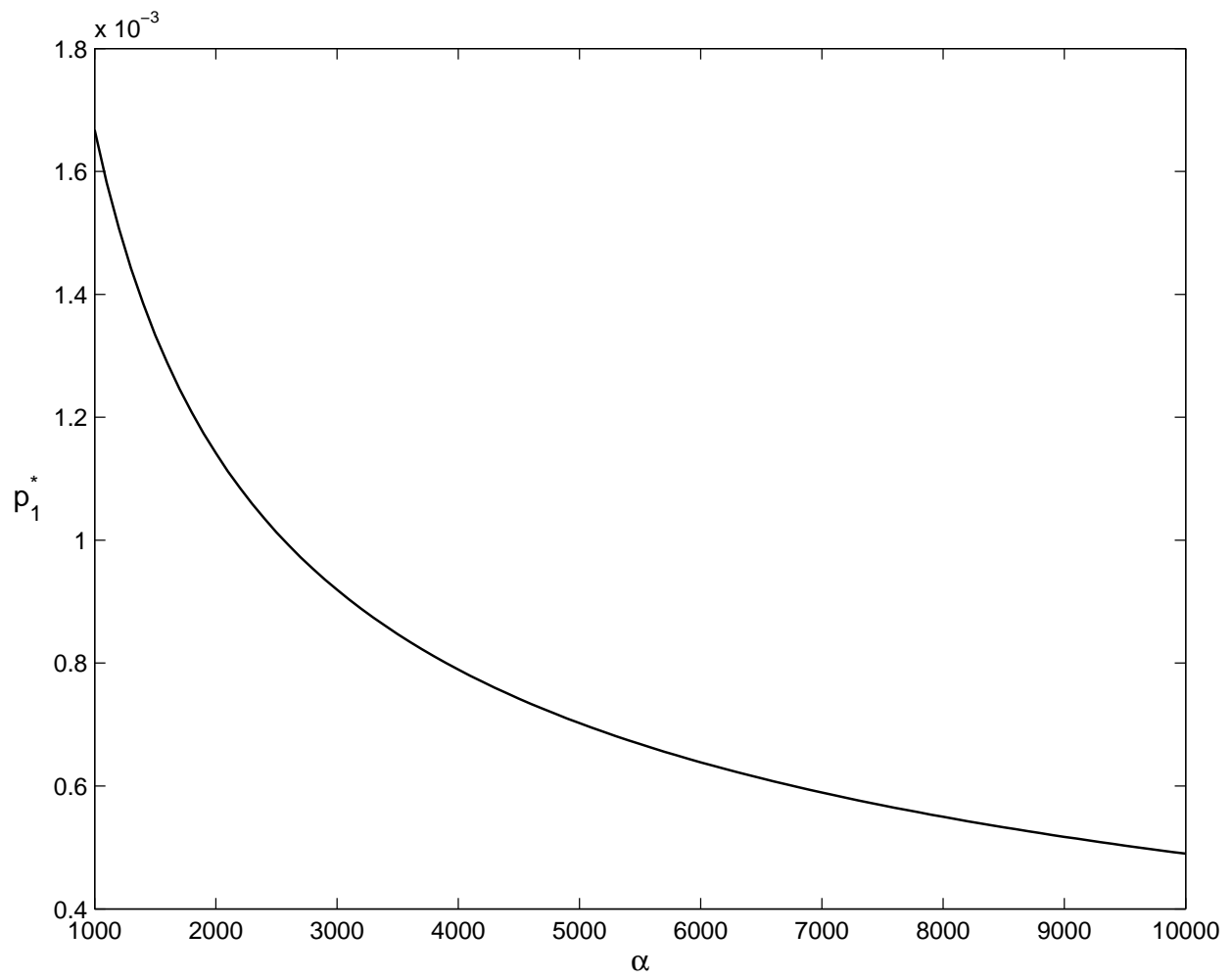
Figure 5: The value of p_1^* as a function of α, assuming that the constraint $p_1 \leq p_0$ is not binding

6   Discussion

We have developed a model to determine the cheapest mix of screening and contact tracing for control of an endemic disease. We considered three cases: 1) the population is large enough (and screening and contact tracing identify a small enough number of cases) such that disease prevalence is unchanged; 2) disease prevalence is reduced and the goal is to minimize long-term cost of control to achieve a specified prevalence; 3) disease prevalence is reduced and the goal is to minimize the total cost of identifying and treating disease plus the cost of untreated disease.

In all three cases, contact tracing is optimal only when disease prevalence is below a threshold value. When prevalence is above the threshold, enough disease cases are found per person screened that the cost per case identified via screening is less than the cost per case identified via contact tracing. When prevalence is below the threshold, mass screening becomes relatively more expensive per case identified compared to contact tracing, and contact tracing becomes part of the minimum cost solution. Additionally, when prevalence increases above the threshold and contact tracing is discontinued, the optimal of level of screening increases discontinuously to substitute for the contact tracing and thus maintain the desired degree of epidemic control.

We can apply our framework (in particular equation (12) or (20)) using data for several infectious diseases to estimate an implied value for the contact tracing threshold, P'. For simplicity we assume perfect screening: $s_{\text{tp}} = 1$ and $s_{\text{fp}} = 0$. For HIV we estimate $C_S = \$2.50$ (Sanders et al., 2005), $C_t = \$175,000$ (Varghese et al., 1999), $C_T = K_T(C_t + \$439)$ (Varghese et al., 1999), and $K_T = 0.12$ (Varghese et al., 1999; Golden, 2002). These values imply a threshold prevalence of $P' = 0.6\%$. In the US, overall HIV prevalence is less than 0.4% (CDC, 2003), so our analysis suggests that contact tracing for HIV in the US makes sense except in high-prevalence areas. For syphilis we estimate $C_S = \$2.88$ (Oxman and Doyle, 1996), $C_t = \$9.50$ (Howell et al., 1997), $C_T = \$83.51$ (Oxman and Doyle, 1996; Howell et al., 1997), and $K_T = 1$ (Oxman and Doyle, 1996), implying a threshold prevalence of $P' = 4\%$. This is well above the prevalence one might expect (e.g., $\sim 0.05\%$, (CDC, 1993a)) suggesting that contact tracing for syphilis in the US makes economic sense. For TB, we estimate $C_S = \$5$, $C_t = \$10,000$ (Rose, 1998), $C_T = \$200 + K_T C_t$ (Marchand et al., 1999), and $K_T = 3$ (Dasgupta et al., 2000), implying a threshold prevalence for contact tracing of $P' = 8\%$. Again, this suggests that TB contact tracing policies in the US make sense. Following similar

reasoning, our analyses indicate that for diseases where the prevalence is high — such as TB in India or HIV in sub-Saharan Africa — contact tracing does not make economic sense.

In the first two formulations we presented, equilibrium disease prevalence is specified exogenously, whereas in the third formulation, the cost-minimizing equilibrium disease prevalence is found as part of the solution to the optimal control problem. This latter formulation has the advantage that it finds the best policy that satisfies the cost-effectiveness criterion implicitly given by the parameter α: the optimal level of disease control is such that the marginal cost of incremental disease control (the cost to find and treat additional disease cases) equals the marginal cost of non-control (the cost associated with undetected disease cases). Such a formulation could be particularly helpful to public health officials who may have to allocate a fixed total amount of funding across programs for the control of a number of infectious diseases. Use of the cost-effectiveness parameter α ensures that the amount of money spent to control a given disease (as determined by solving the optimal control problem) is commensurate with the value of controlling that disease.

Our work has several limitations. We based our analysis on a very simple compartmental model with random mixing. If the infectious disease is known to spread non-randomly (for example, through a contact network), then contact tracing could be more cost effective than assumed in our analysis. If it is possible to identify high-risk individuals or individuals who are centrally located in the contact network, then tracing the contacts of those individuals will (on average) cost less per case identified than will tracing the contacts of randomly selected index cases.

We assumed that the costs of screening and contact tracing were linear in the number of individuals screened and contacts traced. If these costs are nonlinear, such cost functions could readily be incorporated into our analysis. We assumed no delays in identifying new cases via walk-ins, screening, and contact tracing, and we did not include demographic or risk factors such as age, gender, or risk behaviors. For contact tracing, our analysis considered only the expected number of new cases identified per index case. Further research could explore the robustness of our results with more sophisticated models that (partially) lift these limitations, perhaps using simulation and network modeling.

Although simple, our analysis provides several insights. First, contact tracing is only cost effective when prevalence is below a threshold value. This threshold depends on the relative cost per case found by screening versus contact tracing. Second, for a given contact tracing policy (e.g., δ

= 1), the screening rate needed to achieve a given prevalence or identify a specified number of cases is a decreasing function of disease prevalence. As prevalence increases above the threshold (and contact tracing is discontinued), the screening rate jumps discontinuously to a higher level. Third, these qualitative results hold when we consider unchanged or changed prevalence, and short-term or long-term costs.

References

N. T. J. Bailey. *The mathematical theory of infectious diseases and its applications*. Griffin, 1975.

CDC. Primary and secondary syphilis – reported cases and rates in selected cities of >200,000 population: United states and outlying areas, 1989-1993. 1993a. URL `http://wonder.cdc.gov/wonder/STD/OSTD3233/Table_27.html`.

CDC. Tuberculosis control laws — United States, 1993. *MMWR*, 42(RR-15), 1993b. URL `http://www.cdc.gov/mmwr/PDF/rr/rr4215.pdf`.

CDC. Prevention and control of tuberculosis in correctional facilities. *MMWR*, 45(RR-8), 1996. URL `http://www.cdc.gov/mmwr/PDF/rr/rr4508.pdf`.

CDC. Chapter 10 - community TB control: Identification of persons who have clinically active TB. 2000. URL `http://www.cdc.gov/nchstp/tb/pubs/corecurr/Chapter10/Chapter_10_Identification.htm`.

CDC. Sexually transmitted diseases treatment guidelines 2002. 2002. URL `http://www.cdc.gov/STD/treatment/1-2002TG.htm`.

CDC. HIV/AIDS surveillance report. 2003. URL `http://www.cdc.gov/hiv/stats/2003surveillancereport.pdf`.

J. Clarke. Contact tracing for chlamydia: data on effectiveness. *International Journal of STD & AIDS*, pages 187–191, 1998.

D. A. Cohen, S.-Y. Wu, and T. A. Farley. Comparing the cost-effectiveness of HIV prevention

interventions. *Journal of Acquired Immune Deficiency Syndromes*, 37(3), 2004. URL `http://www.jaids.org/pt/re/jaids/fulltext.00126334-200411010-00009.htm`.

F. M. Cowan, R. French, and A. M. Johnson. The role and effectiveness of partner notification in STD control: a review. *Genitourinary Medicine*, 72:247–252, 1996.

K. Dasgupta, K. Schwartzman, R. Marchand, T. N. Tennenbaum, P. Brassard, and D. Menzies. Comparison of cost-effectiveness of Tuberculosis screening of close contacts and foreign-born populations. *American J. Respir. Crit. Care Med.*, 162(6):2079–2086, 2000. URL `http://ajrccm.atsjournals.org/cgi/content/abstract/162/6/2079`.

M. Gold, J. Siegel, L. Russell, and M. Weinstein. *Cost-effectiveness in health and medicine*. Oxford University Press, 1996.

M. R. Golden. HIV partner notification: A neglected prevention intervention. *Sexually Transmitted Diseases*, 29(8):472–475, 2002.

H. W. Hethcote and J. A. Yorke. *Gonorrhea Transmission Dynamics and Control*, volume 56 of *Lecture Notes in Biomathematics*. Springer, 1984. URL `http://biotech.law.lsu.edu/cphl/Models/gon`.

M. R. Howell, W. J. Kassler, and A. Haddix. Partner notification to prevent pelvic inflammatory disease in women: Cost-effectiveness of two strategies. *Sexually Transmitted Diseases*, 24(5):287–292, 1997.

R. Huerta and L. S. Tsimring. Contact tracing and epidemics control in social networks. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 66(5):056115, 2002. URL `http://link.aps.org/abstract/PRE/v66/e056115`.

D. Hutton, D. Tan, S. So, and M. L. Brandeau. Cost effectiveness of Hepatitis B screening and vaccination among adult Asian Pacific Islanders in the United States. Technical report, Stanford University, 2007.

M. Kretzschmar, S. van den Hof, J. Wallinga, and J. van Wijngaarden. Ring vaccination and small-pox control. *Emerging Infectious Diseases*, 10(5), 2004. URL `http://www.cdc.gov/ncidod/eid/vol10no5/pdfs/03-0419.pdf`.

C. R. Macintyre, A. J. Plant, and D. Hendrie. The cost-effectiveness of evidence-based guidelines and practice for screening and prevention of tuberculosis. *Health Economics*, 9(5):411–421, 2000. URL `http://www3.interscience.wiley.com/cgi-bin/abstract/72511803/ABSTRACT`.

R. Marchand, P. Tousignant, and H. Chang. Cost-effectiveness of screening compared to case-finding approaches to tuberculosis in long-term care facilities for the elderly. *Int. J. Epidemiol.*, 28(3):563–570, 1999. URL `http://ije.oupjournals.org/cgi/content/abstract/28/3/563`.

J. Müller, M. Kretzschmar, and K. Dietz. Contact tracing in stochastic and deterministic epidemic models. *Mathematical Biosciences*, 164, 2000.

D. K. Owens. Interpretation of cost-effectiveness analyses. *J Gen Intern Med*, 13:716–717, 1998.

D. K. Owens, R. F. Nease, and R. A. Harris. Cost-effectiveness of human immunodeficiency virus (HIV) screening in acute-care settings. *Archives of Internal Medicine*, 156:394–404, 1996.

G. L. Oxman and L. Doyle. A comparison of the case-finding effectiveness and average costs of screening and partner notification. *Sexually Transmitted Diseases*, 1996.

L. N. Pealer and T. A. Peterman. When it comes to contact notification, HIV is not TB. *The International Journal of Tuberculosis and Lung Disease*, 7:S337–S341, 2003. URL `http://www.ingentaconnect.com/content/iuatld/ijtld/2003/00000007/a00312s3/art00004`.

T. C. Porco, K. A. Holbrook, S. E. Fernyak, D. L. Portnoy, R. Reiter, and T. J. Aragn. Logistics of community smallpox control through contact tracing and ring vaccination: a stochastic network model. *BMC Public Health*, 4(34), 2004. URL `http://www.biomedcentral.com/1471-2458/4/34`.

D. N. Rose. Short-course prophylaxis against Tuberculosis in HIV-infected persons: A decision and cost-effectiveness analysis. *Ann Intern Med*, 129(10):779–786, 1998. URL `http://www.annals.org/cgi/content/abstract/129/10/779`.

P. P. Rubió. Critical value of prevalence for vaccination programmes. the case of Hepatitis A vaccination in Spain. *Vaccine*, 15:1445–1450, 1997.

G. D. Sanders, A. M. Bayoumi, V. Sundaram, S. P. Bilir, C. P. Neukermans, C. E. Rydzak, L. R. Douglass, L. C. Lazzeroni, M. Holodniy, and D. K. Owens. Cost-effectiveness of screening for HIV in the era of highly active antiretroviral therapy. *New England Journal of Medicine*, 352(6):570–585, 2005. URL `http://content.nejm.org/cgi/content/abstract/352/6/570`.

B. Varghese, T. A. Peterman, and D. R. Holtgrave. Cost-effectiveness of counseling and testing and partner notification: a decision analysis. *AIDS*, 13(13):1745–1751, 1999.

Appendix

Here we justify our claim that p_1^* is a nonincreasing function of $\bar{\lambda}$.

Lemma. *Consider control problem A where $\bar{\lambda} = \bar{\lambda}_A$ and problem B where $\bar{\lambda} = \bar{\lambda}_B \leq \bar{\lambda}_A$. Then there exists solutions $p_A^*(t)$ and $p_B^*(t)$ ending up at p_{1A}^* and p_{1B}^* respectively such that $p_{1A}^* \leq p_{1B}^*$.*

Proof. Suppose by contradiction there exist two control problems A and B where $p_{1A}^* > p_{1B}^*$. Now modify the control problems so that $p_0 = p_{1A}^*$. Then $p_A^*(t) = p_{1A}^*$ is an optimal solution of problem A due to the autonomous nature of the control problem. Since A is a relaxation of B, the cost of $p_A^*(\cdot)$ is (weakly) less than the optimal cost of problem B. Hence $p_B^*(t) = p_A^*(t) = p_{1A}^*$ is also an optimal solution of problem B. In this solution $p_{1A} = p_{1B}^*$, a contradiction. $\qquad\square$

We now derive equation (28). Let $\Lambda(t) = \int_0^t \lambda(s)\,\mathrm{d}s$. We perform a change of variables to write p as a function of Λ:

$$p'(\Lambda) = \dot{p}\frac{\mathrm{d}t}{\mathrm{d}\Lambda} = f(p, \frac{\mathrm{d}\Lambda}{\mathrm{d}t}, \delta)\frac{\mathrm{d}t}{\mathrm{d}\Lambda} \tag{29}$$

$$= (1-p)[\beta p + \eta](1 - \gamma(1 + \delta K_T))\frac{\mathrm{d}t}{\mathrm{d}\Lambda} - (1 + \delta K_T)s_{\mathrm{tp}}p. \tag{30}$$

For massive screening in very short times, $\Lambda(t)$ is discontinuous and hence $\lambda(t)$ is a Dirac delta function and $\frac{\mathrm{d}t}{\mathrm{d}\Lambda} = 0$ so

$$p'(\Lambda) = -(1 + \delta K_T)s_{\mathrm{tp}}p(\Lambda) \tag{31}$$

$$p(\Lambda) = p_0 e^{-\Lambda s_{\mathrm{tp}}(1 + \delta K_T)} \tag{32}$$

where $p(0) = p_0$. Hence the number screened to reach prevalence p_1 solves $p(\Lambda) = p_1$:

$$\Lambda = \frac{\log p_0 - \log p_1}{(1 + \delta K_T)s_{\mathrm{tp}}}.$$ (33)

For costs we have a similar differential equation,

$$C'(\Lambda) = C_S + C_t s_{\mathrm{tp}} p(\Lambda) + C_{\mathrm{fS}} s_{\mathrm{fp}}(1 - p(\Lambda)) + \delta C_T s_{\mathrm{tp}} p(\Lambda).$$ (34)

We substitute (32) and then solve the differential equation with the initial condition $C(0) = 0$:

$$C(\Lambda) = \frac{C_t s_{\mathrm{tp}} - C_{\mathrm{fS}} s_{\mathrm{fp}} + \delta s_{\mathrm{tp}} C_T}{(1 + \delta K_T)s_{\mathrm{tp}}} p_0(1 - e^{-\Lambda s_{\mathrm{tp}}(1+\delta K_T)}) + (C_S + C_{\mathrm{fS}} s_{\mathrm{fp}})\Lambda.$$ (35)

Substituting in (33), we arrive at our conclusion.