Online Appendix for "Behaviorally-informed Strategies for Infectious Disease Mitigation: Balancing Incentives and Screening"

Sait Tunç
Grado Department of Industrial & Systems Engineering, Virginia Tech, Blacksburg, VA 24061, saittunc@vt.edu

Berk Görgülü
DeGroote School of Business, McMaster University, Hamilton, ON L8S 4E8, Canada, gorgulub@mcmaster.ca

Marie J. Rabil
Grado Department of Industrial & Systems Engineering, Virginia Tech, Blacksburg, VA 24061, mjrabil20@gmail.com

Douglas R. Bish
Culverhouse College of Business, The University of Alabama, Tuscaloosa, AL 35487, drbish@ua.edu

Ebru K. Bish
Culverhouse College of Business, The University of Alabama, Tuscaloosa, AL 35487, ekbish@ua.edu

Appendix A: Preliminaries

A.1. Summary of Notation

We use upper-case letters for random variables, lower-case letters for their realization, and ";" for probabilistic conditioning, or to indicate dependence on a given decision. For random variable $\mathcal{X}$, $f_{\mathcal{X}}(\cdot)$, $F_{\mathcal{X}}(\cdot)$, $\overline{F}_{\mathcal{X}}(\cdot)$, $F_{\mathcal{X}}^{-1}(\cdot)$, and $\mathcal{S}(\mathcal{X})$ resp. denote the pdf, CDF, tail function, inverse CDF, and sample space.

Table C1: Notation used in the paper

Notation	Description
	Decision variables
d	Vaccination incentive offered to each vaccinated individual
s	Screening rate applied to each unvaccinated individual
s'	Screening rate applied to each unvaccinated individual, updated at time t' in **QM-3**
$^*, J, Q, Q', N$	Superscripts denoting an optimal solution, **JM**, **QM**-2, and **QM**-3, and Nature's parameters, respectively
	Random variables and associated functions
W	A random individual's disutility toward vaccination, with sample space $\mathcal{S}(W) = (-\infty, +\infty)$ and realization w
X_v	Initial vaccination coverage in the population (corresponding to $d = 0,\ s = 0$), with sample space $\mathcal{S}(X_v) = [0, 1]$ and realization x_v
$H(s)$	Population-level disutility of undergoing screening at rate $s \geq 0$, with sample space $\mathcal{S}(H(\cdot))$ containing all functional realizations $h(\cdot)$ satisfying Assumption 1
$\Delta(d, s, X_v, H)$	Induced vaccination coverage resulting from incentive d and screening rate s under uncertainty in initial coverage X_v and screening disutility function $H(\cdot)$, with realization $\delta(d, s; x_v, h)$
$\Delta_v(\cdot),\ \Delta_s(\cdot)$	In the sequential models, the incentive-induced and screening-induced components of induced vaccination coverage, with realizations $\delta_v(\cdot; x_v)$ and $\delta_s(\cdot; x_v, h)$, respectively
$\mathcal{G}(d, s, X_v, H)$	Health reward of interest, where higher values indicate better outcomes
	Parameters and deterministic functions
B	Per-person budget available to cover vaccination incentives for the vaccinated and routine screening expenses for the unvaccinated over the semester
$\gamma > 0$	Cost per screening test
$\overline{S}$	Maximum feasible screening rate per person per semester
$\alpha \in [0, 1]$	Minimum required budget-compliance probability for the mitigation plan

A.2. Definitions

We provide the first- and second-order stochastic dominance definitions (hereafter referred to as usual stochastic order and increasing concave order, respectively (resp.)) used in our analysis.

DEFINITION A.1. (Müller and Stoyan 2002) Let X and X' be two random variables with sample space $\mathcal{S}$ and finite means.

1. We say that X' is smaller than X with respect to (wrt) the usual stochastic order (first-order stochastic dominance), written $X' \leq_{st} X$, if $F_{X'}(x) \geq F_X(x), \forall x \in \mathcal{S}$.
2. We say that X' is smaller than X in increasing concave order (second-order stochastic dominance), written $X' \leq_{icv} X$, if $E[q(X')] \leq E[q(X)]$ for all increasing concave functions $q(.)$ such that expectations exist, equivalently, if and only if $\int_0^x [F_{X'}(t) - F_X(t)]\, dt \geq 0, \forall x \in \mathcal{S}$.

REMARK A.1. (Müller and Stoyan 2002)

(a) $X' \leq_{st} X \Rightarrow E(X') \leq E(X)$.

(b) For any nondecreasing function $u(.) : \mathbb{R} \to \mathbb{R}$, $X' \leq_{st} X \Rightarrow E(u(X')) \leq E(u(X))$.

A.3. Preliminary Derivations and Expressions

In this section, we provide some derivations, expressions, and observations used throughout our analysis. First, recall from (9) that

$$\begin{aligned}\psi(d,s;x_v) &= d\cdot E_H[x_v+\delta(d,s,H;x_v)]+\gamma s\cdot E_H[1-x_v-\delta(d,s,H;x_v)]-B\\ &= d\cdot E_H[F_{W;x_v}\big(d+H(s)\big)]+\gamma s\cdot E_H[\overline{F}_{W;x_v}\big(d+H(s)\big)]-B\\ &= (d-\gamma s)\cdot E_H[F_{W;x_v}\big(d+H(s)\big)]-(B-\gamma s).\end{aligned}$$

Let $\psi_2(s(d,x_v);d,x_v) :=$ LHS - RHS of (11b), that is,

$$\begin{aligned}\psi_2(s(d,x_v);d,x_v) =& d\cdot E_H[\delta_s(s(d,x_v),H;d,x_v)]+\gamma s(d,x_v)\cdot E_H[1-x_v-\delta_s(s(d,x_v);d,x_v)-\delta_v(d;x_v)]\\ &-\Big(B-d\cdot[x_v+\delta_v(d;x_v)]\Big)^+\\ =& d\cdot E_H[F_{W;x_v}(d+H(s))]+\gamma s(d,x_v)\cdot E_H[\overline{F}_{W;x_v}\big(d+H(s)\big)]-d\cdot F_{W;x_v}(d)\\ &-\Big(B-d\cdot F_{W;x_v}(d)\Big)^+.\end{aligned}$$

Notice that $\frac{\partial}{\partial s}\psi(d,s;x_v) = \frac{\partial}{\partial s}\psi_2((s(d,x_v);d,x_v)$, where

$$\begin{aligned}\psi'(s) := \frac{\partial}{\partial s}\psi(d,s;x_v) &= (d-\gamma s)\cdot E_H[f_{W;x_v}(d+H(s))\cdot H'(s)]+\gamma\cdot[1-E_H[F_{W;x_v}\big(d+H(s)\big)]\,]\\ &= (d-\gamma s)\cdot E_H[f_{W;x_v}(d+H(s))\cdot H'(s)]+\gamma\cdot E_H[1-F_{W;x_v}\big(d+H(s)\big)]\\ &= (d-\gamma s)\cdot E_H[f_{W;x_v}(d+H(s))\cdot H'(s)]+\gamma\cdot E_H[\overline{F}_{W;x_v}\big(d+H(s)\big)]. \qquad \text{(S.1)}\end{aligned}$$

Using (S.1), we deduce the following behavior for $\psi'(s)$ under different conditions:

$$\psi'(s)\leq 0 \quad \text{iff } s\geq d/\gamma+\frac{E_H\left[\overline{F}_{W;x_v}\big(d+H(s)\big)\right]}{E_H[f_{W;x_v}(d+H(s))\cdot H'(s)]} \qquad \text{(S.2a)}$$

$$\psi'(s)>0 \quad \text{otherwise.} \qquad \text{(S.2b)}$$

REMARK A.2. (a) By Assumption 3, the optimal updated screening rate s'^* (the solution to **QM-3** deterministic third stage) is uniquely determined for each given $(d, x_v, s, h(\cdot))$: If there is remaining budget at time t', s'^* is set to the maximum screening rate that fully utilizes the remaining budget, provided it does not exceed $\overline{S}$; otherwise, it is set to $\overline{S}$. If the remaining budget at time t' is zero, $s'^* = 0$.

(b) Using the characterization for s'^* in part (a), stages 2 and 3 of **QM-3** can be consolidated into a stage 2 formulation, which is analogous to **QM-2**, with the only distinction being the expected cost function.

This expectation, taken over the random function $H(\cdot)$, now reflects the combined costs of stages 2 and 3, that is, the expected value of the following function:

$$\begin{aligned} C_2(s(d,x_v),H;d,x_v) :=& d\cdot\delta_s(s(d,x_v),H;d,x_v)+[1-x_v-\delta_s(s(d,x_v),H;d,x_v)-\delta_v(d;x_v)]\cdot\gamma \\ &\cdot\left[\frac{t'}{T}\cdot s(d,x_v)+\frac{(T-t')}{T}\cdot s'(s(d,x_v),H)\right]+\left[d-\gamma\cdot\frac{(T-t')}{T}\cdot s'(s(d,x_v),H)\right] \\ &\cdot\delta_s'\Big(s'(s(d,x_v),H),s(d,x_v),H;d,x_v\Big), \end{aligned} \tag{S.3}$$

where $\delta_s'(s',s,H;d,x_v)$ is defined in Eq. (7) and $s'(s(d,x_v),H)$ can be uniquely determined given s,d,x_v,h, as discussed in part (a). In (S.3), the first component represents incentive payments for vaccinations induced the initial screening rate s at semester onset. The second component captures screening expenses without accounting for additional vaccinations that may occur at time t' due to the updated rate s' (which is positive only if $s'>s$, see Eq. (7)). The third component adjusts these screening expenses from the second term to reflect new vaccinations (if any) at time t', and is positive only if $s'>s$. Specifically, the newly vaccinated (proportion $\delta_s'(s',s,H;d,x_v)$) receive the incentive payment d instead of incurring screening costs between t' and T, which would otherwise amount to $\gamma\cdot\frac{(T-t')}{T}\cdot s'(s(d,x_v),H)$.

Based on Remark A.2, we extend the budget deficit function $\psi(\cdot)$ to **QM-3** by accounting for expenditures over all stages. For a given realization $x_v\in\mathcal{S}(X_v)$, the total deficit (surplus) for decision tuple (d,s,s') is given by

$$\psi(d,s,s';x_v)=d\cdot(x_v+\delta_v(d;x_v))+E_H[C_2(s(d,x_v),H;d,x_v)]-B.$$

In this setting, we define the set of roots of the budget deficit , $\psi(\cdot;x_v)$, conditional on realization x_v, wrt s' as follows. For any $x_v\in\mathcal{S}(X_v)$, let $s'_{root}(d,s;x_v):=\{s'\geq 0:\psi(d,s,s';x_v)=0\},\forall d\geq 0,s\geq 0$ (s'-root for a given d and s), when exists. When the set $s'_{root}(d,s;x_v)$ is nonempty, to construct the Pareto set for **QM-3**, it suffices to consider the largest feasible s'-root, denoted by $s'_{\max}(d,s;x_v):=\max\{s'_{root}(d,s;x_v):s'_{root}(d,s;x_v)\leq\overline{S}\},\forall d\geq 0,s\geq 0$, which captures the most aggressive mid-semester screening adjustment that utilizes all remaining budget.

A.4. Preliminary Structural Properties

The following property follows by the definition of the s- and d-roots.

PROPERTY A.1. *For any $x_v\in\mathcal{S}(X_v)$:*

(a) *$d_{root}(0;x_v)\geq B$, and uniquely solves $d_{root}(0;x_v)\cdot F_{W;x_v}\big(d_{root}(0;x_v)\big)=B$.*

(b) *$s_{root}(d;x_v)$ is not necessarily unique unless $d=B$. Any $s_{root}(d;x_v)$, provided it exists, satisfies:*

$$s_{root}(d;x_v)=\begin{cases}<{}^{B}\!/_{\gamma}, & \forall d:B<d\leq d_{root}(0;x_v)\\ >{}^{B}\!/_{\gamma}, & \forall d:0\leq d<B\\ ={}^{B}\!/_{\gamma}, & \textit{for } d=B\\ \textit{does not exist}, & \forall d:d>d_{root}(0;x_v)\end{cases}.$$

Proof of Property A.1.

(a) We have $\psi(d,0;x_v) := d \cdot F_{W;x_v}\big(d\big) - B = 0$ and $\frac{\partial \psi(d,0;x_v)}{\partial d} = F_{W;x_v}\big(d\big) + d \cdot f_{W;x_v}\big(d\big) > 0, \forall d > 0$. Since $\psi(d,0;x_v) = -B < 0$ when $d = 0$ and $\lim_{d\to\infty} \psi(d,0;x_v) = \infty$, by the intermediate value theorem, there exists only one $d \in (0,\infty)$ such that $\psi(d,0;x_v) = 0$. This unique d-root trivially satisfies $d \geq d \cdot F_{W;x_v}\big(d\big) = B \geq 0$.

(b) $s_{root}(d;x_v)$ is not necessarily unique unless $d = B$, since only for $d = B$, $\psi(d,s;x_v) = 0$ simplifies to $E_H\left[F_{W;x_v}\big(d + h(s)\big)\right] = 1$, leading to a closed-form solution for d.

Next, by definition of $s_{root}(d;x_v)$ and reorganizing $\psi(d,s;x_v) = 0$ as $(d - \gamma s) \cdot E_H\left[F_{W;x_v}\big(d + H(s)\big)\right] = B - \gamma s$, we obtain the following:

- $\forall d : B < d \leq d_{root}(0;x_v)$, we have

$$\begin{aligned}\gamma s \cdot E_H\left[\overline{F}_{W;x_v}\big(d + H(s)\big)\right] &= B - d \cdot E_H\left[F_{W;x_v}\big(d + H(s)\big)\right] \\ &< d - d \cdot E_H\left[F_{W;x_v}\big(d + H(s)\big)\right] \\ &= d \cdot E_H\left[\overline{F}_{W;x_v}\big(d + H(s)\big)\right].\end{aligned}$$

  Therefore, $\forall d : B < d \leq d_{root}(0;x_v)$, $\gamma s < d$. Because $(d - \gamma s) > 0$ and $E_H\left[F_{W;x_v}\big(d + H(s)\big)\right] > 0$, we have $B - \gamma s > 0$, i.e., $B > \gamma s$. Hence, $\forall d : B < d \leq d_{root}(0;x_v)$, we have $s_{root}(d;x_v) < \frac{B}{\gamma}$.

- $\forall d : 0 \leq d < B$, we have

$$\begin{aligned}\gamma s \cdot E_H\left[\overline{F}_{W;x_v}\big(d + H(s)\big)\right] &= B - d \cdot E_H\left[F_{W;x_v}\big(d + H(s)\big)\right] \\ &> d - d \cdot E_H\left[F_{W;x_v}\big(d + H(s)\big)\right] \\ &= d \cdot E_H\left[\overline{F}_{W;x_v}\big(d + H(s)\big)\right].\end{aligned}$$

  Therefore, $\forall d : 0 \leq d < B$, $\gamma s > d$. Because $(d - \gamma s) < 0$ and $E_H\left[F_{W;x_v}\big(d + H(s)\big)\right] > 0$, we have $B - \gamma s < 0$, i.e., $B < \gamma s$. Hence, $\forall d : 0 \leq d < B$, we have $s_{root}(d;x_v) > \frac{B}{\gamma}$.

- For $d = B$, we have $B \cdot E_H\left[F_{W;x_v}\big(B + H(s)\big)\right] + \gamma s \cdot E_H\left[\overline{F}_{W;x_v}\big(B + H(s)\big)\right] = B$, and it is easy to verify that $s_{root}(d;x_v) = \frac{B}{\gamma}$ is a solution. Further, it is the only solution because we cannot have $s_{root}(d;x_v) > \frac{B}{\gamma}$ (since $d \not< B$) neither $s_{root}(d;x_v) < \frac{B}{\gamma}$ (since $d \not> B$).

- $\forall d : d > d_{root}(0;x_v)$, we have $s < 0$ and the budget is already exhausted by the monetary incentive ($\psi(d,s;x_v) > 0$), hence such d values are not feasible, and $s_{root}(d;x_v)$ does not exist.

□

We next establish key structural properties of the budget deficit function $\psi(\cdot)$.

PROPERTY A.2. (a) *The budget deficit function, $\psi(d,s;x_v)$, given in (9), is increasing in x_v if $d > \gamma s$; decreasing in x_v if $d < \gamma s$; and becomes independent of x_v if $d = \gamma s$. Then:*

(a-i) *For any* $(d,s): d > \gamma s$, *we define* $x_{max}(d,s) := argmax_{x_v \in \mathcal{S}(X_v)}\{\psi(d,s;x_v) \le 0\}$. *If no such* $x_v \in \mathcal{S}(X_v)$ *exists, then we set* $x_{max}(d,s) = -\infty$. *Equivalently, we have* $x_{max}(d,s)$: $Pr\big(\Psi(d,s,X_v) \le 0\big) = F_{X_v}(x_{max}(d,s))$.

(a-ii) *For any* $(d,s): d < \gamma s$, *we define* $x_{min}(d,s) := argmin_{x_v \in \mathcal{S}(X_v)}\{\psi(d,s;x_v) \le 0\}$. *If no such* $x_v \in \mathcal{S}(X_v)$ *exists, then we set* $x_{min}(d,s) = +\infty$. *Equivalently, we have* $x_{min}(d,s)$: $Pr\big(\Psi(d,s,X_v) \le 0\big) = \overline{F}_{X_v}(x_{min}(d,s))$.

(b) *Consider any* $x_v \in \mathcal{S}(X_v)$. *We have that* $\psi(d,0;x_v) \le 0$ *for all* $d \le d_{root}(0;x_v)$; *and* $\psi(d,0;x_v) > 0$, *otherwise. Further:*

(b-i) $\psi(d,s;x_v)$ *is increasing in* $s \in [0, \frac{d}{\gamma}]$;

(b-ii) $\lim_{s\to\infty}\big\{\gamma s \cdot \overline{F}_{W;x_v}\big(d + h(s)\big)\big\} \to \infty$, *hence* $\lim_{s\to\infty}\psi(d,s;x_v) \to \infty,\ \forall d \ge 0$.

Proof of Property A.2.

(a) The derivative of $\psi(d,s;x_v) := (d-\gamma s)\cdot E_H\left[F_{W;x_v}\big(d+H(s)\big)\right] - (B-\gamma s)$ wrt x_v is given by

$$\frac{\partial \psi(d,s;x_v)}{\partial x_v} = (d-\gamma s)\cdot\frac{\partial E_H\left[F_{W;x_v}\big(d+H(s)\big)\right]}{\partial x_v} = (d-\gamma s)\cdot E_H\left[\frac{\partial F_{W;x_v}\big(d+H(s)\big)}{\partial x_v}\right] \quad \text{(S.4)}$$

where the interchanging of expectation and differentiation follows due to the dominated convergence theorem; also $E_H\left[\frac{\partial F_{W;x_v}\left(d+h(s)\right)}{\partial x_v}\right] > 0$ by Assumption 2. Therefore, we have $\frac{\partial \psi(d,s;x_v)}{\partial x_v} > 0$, for $(d,s): d > \gamma s$, and, $\frac{\partial \psi(d,s;x_v)}{\partial x_v} < 0$, for $(d,s): d < \gamma s$, for any given x_v. If $d = \gamma s$, on the other hand, $\psi(d,s;x_v) = \gamma s - B$, which is a constant that is independent of x_v.

(a-i) For any $(d,s): d > \gamma s$, recall from (8b′) that $Pr\big(\Psi(d,s,X_v) \le 0\big) = \int_{x_v \in \mathcal{S}(X_v)} 1_{\{\psi(d,s;x_v)\le 0\}} \cdot f_{X_v}(x_v)\,dx_v$, which can be rewritten as, $\int_{x_v:\psi(d,s;x_v)\le 0} f_{X_v}(x_v)\,dx_v$ since the uncertainty is on X_v, and d and s are fixed. Thus, by definition of $x_{max}(d,s)$ and because $\psi(d,s;x_v)$ is monotone increasing in x_v $\forall (d,s): d > \gamma s$ (Property A.2(a)), we can show that $x_v \le x_{max}(d,s)$ iff $\psi(d,s;x_v) \le 0$ and $x_v > x_{max}(d,s)$ iff $\psi(d,s;x_v) > 0$. Therefore, $\int_{x_v:\psi(d,s;x_v)\le 0} f_{X_v}(x_v)dx_v$ can be rewritten as $\int_{x_v:x_v \le x_{max}(d,s)} f_{X_v}(x_v)dx_v$ which equals $F_{X_v}(x_{max}(d,s))$.

(a-ii) Similar to the previous part, for any $(d,s): d < \gamma s$, we can rewrite $Pr\big\{\Psi(d,s,X_v) \le 0\big\}$ as $\int_{x_v:\psi(d,s;x_v)\le 0} \cdot f_{X_v}(x_v)dx_v$ since the uncertainty is on X_v, and d and s are fixed. Thus, by definition of $x_{min}(d,s)$, and because $\psi(d,s;x_v)$ is monotone decreasing in x_v $\forall (d,s): d < \gamma s$ (Property A.2(a)), we can show that $x_v \ge x_{min}(d,s)$ iff $\psi(d,s;x_v) \le 0$ and $x_v < x_{min}(d,s)$ iff $\psi(d,s;x_v) > 0$. Therefore, $\int_{x_v:\psi(d,s;x_v)\le 0} f_{X_v}(x_v)dx_v$ can be rewritten as $\int_{x_v:x \ge x_{min}(d,s)} f_{X_v}(x_v)dx_v$ which equals $1 - F_{X_v}(x_{min}(d,s)) = \overline{F}_{X_v}(x_{min}(d,s))$.

(b) Since $\psi(d,0;x_v) := d \cdot F_{W;x_v}\big(d\big) - B$, we have $\psi(d,0;x_v) \le 0$ for all $d \le d_{root}(0;x_v)$, where $d_{root}(0;x_v)$ is the unique solution to $d_{root}(0;x_v)\cdot F_{W;x_v}\big(d_{root}(0;x_v)\big) = B$ (Property A.1(a)).

(b-i) Using (S.2b) (§A.3), we deduce that $\psi'(s) > 0\ \forall s \ge 0 : s < d/\gamma + \frac{E_H\left[\overline{F}_{W;x_v}(d+H(s))\right]}{E_H\left[f_{W;x_v}(d+H(s))\cdot H'(s)\right]}$, where $\frac{E_H\left[\overline{F}_{W;x_v}(d+H(s))\right]}{E_H\left[f_{W;x_v}(d+H(s))\cdot H'(s)\right]} > 0$. Hence, $\psi'(s) > 0\ \forall s \ge 0 : s \le d/\gamma$.

(b-ii) Since every $h(s) \in \mathcal{S}(H(\cdot))$ is bounded by Assumption 1, $\exists\ \overline{\mathrm{H}} < \infty$ such that $h(s) \leq \overline{\mathrm{H}}$ for any s and $h(s) \in \mathcal{S}(H(\cdot))$, and consequently, we have $\overline{F}_{W;x_v}\big(d + h(s)\big) \geq \overline{F}_{W;x_v}\big(d + \overline{\mathrm{H}}\big) > 0$ (because both d and $\overline{\mathrm{H}}$ are finite), which implies $\lim_{s\to\infty}\big\{\gamma s \cdot \overline{F}_{W;x_v}\big(d + h(s)\big)\big\} \geq \lim_{s\to\infty}\big\{\gamma s \cdot \overline{F}_{W;x_v}\big(d + \overline{\mathrm{H}}\big)\big\} \to \infty$ for any $h(s) \in \mathcal{S}(H(\cdot))$. As a result, we have $\lim_{s\to\infty}\big\{\gamma s \cdot E_H\big[\overline{F}_{W;x_v}\big(d + H(s)\big)\big]\big\} \to \infty$, and $\lim_{s\to\infty}\psi(d, s; x_v) \to \infty,\ \forall d \geq 0$.

□

Appendix B: Proofs of Analytical Results

PROPOSITION 1. *Consider* **JM**. *The chance constraint* (8b) *in any optimal solution can be equivalently expressed as the following set of constraints:*

$$\psi(d, s; x^{cr}_{max}) = 0, \quad \forall (d, s) : d_{root}(0; x^{cr}_{max}) \geq d > B > \gamma s, \tag{14a}$$

$$\psi(d, s; x^{cr}_{min}) = 0, \quad \forall (d, s) : 0 \leq d < B < \gamma s, \tag{14b}$$

$$d = \gamma s \leq B, \tag{14c}$$

where $x^{cr}_{max} := F^{-1}_{X_v}(\alpha)$ *and* $x^{cr}_{min} := \overline{F}^{-1}_{X_v}(\alpha)$.

Proof of Proposition 1. First recall from (8b′) that the chance constraint (8b) can be written as $Pr\Big\{\Psi(d, s, X_v) \leq 0\Big\} = \int_{x_v \in \mathcal{S}(X_v)} 1_{\{\psi(d,s;x_v)\leq 0\}} f_{X_v}(x_v)\ dx_v \geq \alpha$. We prove that in an optimal **JM** solution, denoted by (d^*, s^*), Constraint (8b) is tight unless the solution satisfies $d^* = \gamma s^* = B$. Suppose, to the contrary, that the optimal **JM** solution satisfies, $d^* \neq \gamma s^*$ and Constraint (8b) is not tight, i.e., $Pr\{\Psi(d^*, s^*, X_v) \leq 0\} > \alpha$. Since $\psi(d, s; x_v) := (d - \gamma s) \cdot E_H\big[F_{W;x_v}\big(d + H(s)\big)\big] - (B - \gamma s)$, we have

$$\frac{\partial \psi(d, s; x_v)}{\partial d} := E_H\big[F_{W;x_v}\big(d + H(s)\big)\big] + (d - \gamma s) \cdot E_H\big[f_{W;x_v}\big(d + H(s)\big)\big].$$

Hence we have $\frac{\partial}{\partial d}\psi(d, s; x_v) > 0, \forall (d, s) : d > \gamma s, \forall x_v$, and $\lim_{d\to\infty}\psi(d, s; x_v) = \infty$. In addition, $\forall\, d, s, x_v \geq 0, \exists\, \epsilon_1 > 0 : \forall d' > d + \epsilon_1,\ \frac{\partial}{\partial d}\psi(d', s; x_v) > 0$, i.e. $\psi(d', s; x_v)$ is increasing in d'. This implies that random variable $\Psi(d', s, X_v)$ is stochastically increasing in d' and hence, $Pr(\Psi(d', s, X_v) \leq 0)$ is decreasing in d'. Therefore, since $Pr\{\Psi(d^*, s^*, X_v) \leq 0\} > \alpha$ and since the objective function is increasing in d, $\exists\, \epsilon > 0 :$ $\Pr\{\Psi(d^*, s^*, X_v) \leq 0\} > Pr\{\Psi(d^* + \epsilon, s^*, X_v) \leq 0\} = \alpha$. The solution $(d^* + \epsilon, s^*)$ is feasible and yields a higher objective function, contradicting the optimality of (d^*, s^*).

Therefore, the chance constraint (Constraint (8b)) must be tight at (d^*, s^*), that is, $Pr(\Psi(d^*, s^*, X_v) \leq 0) = \alpha$, for $d^* \neq \gamma s^*$. It is also easy to observe that $Pr(\Psi(d^*, s^*, X_v) \leq 0) = 1$ for $d^* = \gamma s^* \leq B$.

Accordingly, if $d^* > \gamma s^*$, using Property A.2(a-i), we have the following:

$$Pr(\Psi(d^*, s^*, X_v) \leq 0) = F_{X_v}(x_{max}(d^*, s^*)) = \alpha;$$

$$\Leftrightarrow x_{max}(d^*, s^*) = F^{-1}_{X_v}(\alpha) = x^{cr}_{max}.$$

Because $\psi(d, s; x_v)$ is monotone increasing in x_v, $\forall d > \gamma s$ (Property A.2(a)), we deduce that $\psi(d^*, s^*; x^{cr}_{max}) = 0$ if $d^* > \gamma s^*$.

Similarly, for $d^* < \gamma s^*$, using Property A.2(a-ii), we have the following:

$$Pr(\Psi(d^*, s^*, X_v) \leq 0) = \overline{F}_{X_v}(x_{min}(d^*, s^*)) = \alpha;$$
$$\Leftrightarrow x_{min}(d^*, s^*) = \overline{F}^{-1}_{X_v}(\alpha) = x^{cr}_{min}.$$

Because $\psi(d, s; x_v)$ is monotone decreasing in x_v, $\forall d < \gamma s$ (Property A.2(a)), we conclude that $\Psi(d^*, s^*; x^{cr}_{min}) = 0$ if $d^* < \gamma s^*$. $\square$

PROPOSITION 2. *Consider* **QM-2**. *An optimal stage 1 solution satisfies that* $d^{*Q} \leq d_{root}(0; x^{cr}_{max})$.
For any $d \leq d_{root}(0; x^{cr}_{max})$ *and realization* $x_v \in \mathcal{S}(X_v)$, *an optimal stage 2 solution* $s^{*Q}(d; x_v)$ *follows:*

(a) *If* $x_v \geq x_{max}(d, 0)$, *then* $\psi(d, 0; x_v) \geq 0$, *hence* $s^{*Q}(d; x_v) = 0$.

(b) *Otherwise* $(x_v < x_{max}(d, 0))$:

(b-i) *If* $\psi(d, \overline{S}; x_v) \leq 0$, *then* $s^{*Q}(d; x_v) = \overline{S}$.

(b-ii) *Otherwise* $\big(\psi(d, \overline{S}; x_v) > 0\big)$, $s^{*Q}(d; x_v) = s_{\max}(d; x_v)$.

Hence, **QM-2** *can be equivalently formulated as a single-stage problem, which optimizes over* d *only:*

$$\max_{0 \leq d \leq d_{root}(0; x^{cr}_{max})} \left\{ \int_{x_v < x_{max}(d,0)} E_H[\mathcal{G}\left(d, s^{*Q}(d; x_v); x_v\right)] f_{X_v}(x_v)\, dx_v + \int_{x_v \geq x_{max}(d,0)} E_H[\mathcal{G}\left(d, 0; x_v\right)] f_{X_v}(x_v)\, dx_v \right\}.$$

Proof of Proposition 2. The result that $d^{*Q} \leq d_{root}(0; x^{cr}_{max})$ follows directly from Property A.2(b), because $\Psi(d, 0; x^{cr}_{max}) > 0$ for any $d > d_{root}(0; x^{cr}_{max})$, that is, the chance constraint in (10b) is not satisfied. Using the definition of $x_{max}(d, 0)$ in Property A.2(a-i), we deduce that $d \geq d_{root}(0; x_v)$ is equivalent to $x_v \geq x_{max}(d, 0)$, and similarly, $d < d_{root}(0; x_v)$ is equivalent to $x_v < x_{max}(d, 0)$.

(a) If $d \geq d_{root}(0; x_v)$, we know from Property A.2(b) that $\psi(d, 0; x_v) \geq 0$, i.e., the budget is depleted in stage 1 and the RHS of (11b), i.e., $\Big(B - d \cdot [x_v + \delta_v(d; x_v)]\Big)^+$, is zero. This leads the LHS of (11b) to be zero for any feasible solution, and as a result, we have $s^{*Q}(d; x_v) = 0$.

(b) Recall that $\psi_2(s; d, x_v) =$ LHS - RHS of (11b) (see §A.3). For $d < d_{root}(0; x_v)$ (which is equivalent to $x_v < x_{max}(d, 0)$), the RHS of (11b) simply reduces to $B - d \cdot [x_v + \delta_v(d; x_v)]$ and we have $\psi_2(s; d, x_v) = \psi(d, s; x_v)$. Thus, the stage 2 feasible set becomes, $s := \{s \in [0, \overline{S}] \mid \psi(d, s; x_v) \leq 0\}$. Recall that $s_{root}(d; x_v) := \{s \geq 0 :\ \psi(d, s; x_v) = 0\}$.

(b-i) If $\psi(d, \overline{S}; x_v) \leq 0$, then by Assumption 3, $s^{*Q}(d; x_v) = \overline{S}$.

(b-ii) First, we prove that if $\psi(d, \overline{S}; x_v) > 0$, then the optimal $s^{*Q}(d; x_v) \in [0, \overline{S})$ satisfies $\psi(d, s^{*Q}(d; x_v); x_v) = 0$. Assume that $\psi(d, s^{*Q}(d; x_v); x_v) \neq 0$. For any feasible s, we have $\psi(d, s; x_v) \leq 0$, which implies that $\psi(d, s^{*Q}(d; x_v); x_v) < 0$. Since $\psi(d, \overline{S}; x_v) > 0$, by the intermediate value theorem, there exists $s' \in (s^{*Q}(d; x_v), \overline{S})$ such that $\psi(d, s'; x_v) = 0$, which gives a contradiction since $s^{*Q}(d; x_v)$ was the maximal value possible. Therefore,

we have $s^{*Q}(d;x_v) = s_{\max}(d;x_v)$ by Assumption 3. Further, since $\psi(d,0;x_v) = d\cdot[x_v + \delta_v(d;x_v)] - B \le 0$ and $\psi(d,\overline{S};x_v) > 0$, using the intermediate value theorem, we conclude that $|s_{root}(d;x_v)| \ge 1$, hence $s_{\max}(d;x_v) \neq \emptyset$.

Finally, using (13) and the characterization of $s^{*Q}(d;x_v)$ given in the proof of Proposition 2, we get

$$E_{X_v}[\mathcal{G}(d,s,X_v)] = \int\limits_{x_v<x_{max}(d,0)} E_H\left[\mathcal{G}\left(d,s^{*Q}(d;x_v),H;x_v\right)\right] f_{X_v}(x_v)\,dx_v + \int\limits_{x_v\ge x_{max}(d,0)} \mathcal{G}(d,0;x_v)\,f_{X_v}(x_v)\,dx_v.$$

In addition, from (8b′), we have that $Pr\{\Psi(d,s,X_v)\le 0\} = \int_{x_v\in\mathcal{S}(X_v)} 1_{\{\psi(d,s;x_v)\le 0\}} f_{X_v}(x_v)\,dx_v \ge \alpha$. We also have from the analysis above that $x_v < x_{max}(d,0)$ if $\psi(d,0;x_v) < 0$ and $x_v \ge x_{max}(d,0)$, otherwise. Hence, the following holds:

$$\int_{x_v\in\mathcal{S}(X_v)} 1_{\{\psi(d,s;x_v)\le 0\}} f_{X_v}(x_v)\,dx_v = \int_{x_v:x_v\le x_{max}(d,0)} f_{X_v}(x_v)dx_v = F_{X_v}(x_{max}(d,0)).$$

As a result, we get that $Pr\left\{\Psi(d,s,X_v)\le 0\right\} \ge \alpha \;\equiv\; F_{X_v}(x_{max}(d,0)) \ge \alpha \;\equiv\; x_{max}(d,0) \ge F^{-1}_{X_v}(\alpha)$. Since $x_{max}(d,0)$ is nonincreasing in d (this holds because $\psi(d,s;x_v)$ is increasing in d and in x_v for $d\ge\gamma s$, Property A.2(a)), the highest feasible d, i.e., $d_{root}(0;x_v)$, is given when $x_{max}(d,0) = F^{-1}_{X_v}(\alpha) = x^{cr}_{max}$. Hence, the chance constraint $Pr\left\{\Psi(d,s,X_v)\le 0\right\} \ge \alpha$ is equivalent to $d \le d_{root}(0;x^{cr}_{max})$. □

COROLLARY 1. *Both* **JM** *and* **QM-2** *Pareto sets span all three strategy classes:*

JM: *incentive-centric strategies*— $(d, s_{\max}(d;x^{cr}_{max})) : d\in(B, d_{root}(0;x^{cr}_{max})]$ *and* $s_{\max}(d;x^{cr}_{max})$ *exists*;
screening-centric strategies— $(d, s_{\max}(d;x^{cr}_{min})) : d\in[0,B)$ *and* $s_{\max}(d;x^{cr}_{min})$ *exists*;
balanced strategy— $(d,s) : d = B$ *and* $s = \dfrac{B}{\gamma} \le \overline{S}$.

QM-2*: incentive-centric if* $d\in(B, d_{root}(0;x^{cr}_{max})]$*; screening-centric if* $d\in[0,B)$*; balanced if* $d = B$.

Proof of Corollary 1.

(a) The result follows from Property A.1(b) and Proposition 1 when the $s\le\overline{S}$ constraint is included.

(b) The result follows from Property A.1(b) and Proposition 2.

□

PROPOSITION 3. *For any* $x_v\in\mathcal{S}(X_v)$ *and* $d\le d_{root}(0;x_v)$, $s_{\max}(d;x_v)$ *behaves as follows:*

(a) *For any incentive-centric strategy (*$B < d \le d_{root}(0;x_v)$*),* $s_{\max}(d;x_v)$ *is decreasing in* d*, with* $s_{\max}(d_{root}(0;x_v);x_v) = 0$.

(b) *For any screening-centric strategy (*$0\le d < B$*),* $s_{\max}(d;x_v)$ *is not necessarily monotone in* d*. However,* $s_{\max}(d;x_v)$ *becomes nonincreasing in* d *if:*

$$\frac{E_H\left[f_{W;x_v}(d+H(s_{\max}(d;x_v)))\right]}{E_H\left[F_{W;x_v}(d+H(s_{\max}(d;x_v)))\right]} \le \frac{1}{\gamma s_{\max}(d;x_v) - d},\ \forall d\in(0,B). \tag{15}$$

Proof of Proposition 3. We use the implicit function theorem and Assumption 4 to study the impact of the monetary award on the optimal screening rate, i.e., $\frac{\partial s_{\max}(d;x_v)}{\partial d}$, for any internal solution $s_{\max}(d;x_v) \in (0,\overline{S})$. Given that $d \le d_{root}(0;x_v)$, we have $\psi(d, s_{\max}(d;x_v);x_v) = 0$ by definition of $s_{\max}(d;x_v) \in (0,\overline{S})$. Therefore, for any interior solution $s_{\max}(d;x_v) \in (0,\overline{S})$, we can derive the implicit derivative as

$$\frac{\partial\, s_{\max}(d;x_v)}{\partial d} = -\frac{\frac{\partial\,\psi(d,s;x_v)}{\partial d};_{s=s_{\max}(d;x_v)}}{\frac{\partial\,\psi(d,s;x_v)}{\partial s};_{s=s_{\max}(d;x_v)}}. \tag{S.5}$$

We first prove by contradiction that the denominator $\frac{\partial\,\psi(d,s;x_v)}{\partial s} > 0$ when $d < d_{root}(0;x_v)$.

From (S.1) (§A.3), we have the following:

$$\frac{\partial\,\psi(d,s;x_v)}{\partial s} = (d-\gamma s)\cdot E_H\left[f_{W;x_v}(d+H(s))\cdot H'(s)\right] + \gamma\cdot E_H\left[\overline{F}_{W;x_v}\big(d+H(s)\big)\right] > 0,$$

for any $d > \gamma s$, or equivalently, $d > B$ (Property A.1(b)).

When $d < B$, we have $d < \gamma s$ (Property A.1(b)). In this case, we show that $\frac{\partial\,\psi(d,s;x_v)}{\partial s} > 0, \forall d < B$ given Assumption 4. Based on the derivation in §A.3, we deduce from (S.2a) that, for any $s \in (0,\overline{S})$ satisfying $s \ge d/\gamma + \frac{E_H\left[\overline{F}_{W;x_v}(d+H(s))\right]}{E_H\left[f_{W;x_v}(d+H(s))\cdot H'(s)\right]}$, $\exists\,\epsilon > 0$ such that $s+\epsilon \in (0,\overline{S})$ and $\psi(d,s;x_v) > \psi(d,s+\epsilon;x_v)$ and therefore s cannot be optimal (because $s+\epsilon$ is feasible with a higher objective value). Therefore, the optimal solution (denoted by s^*) should satisfy $s^* < d/\gamma + \frac{E_H\left[\overline{F}_{W;x_v}(d+H(s^*))\right]}{E_H\left[f_{W;x_v}(d+H(s^*))\cdot H'(s^*)\right]}$, which can be re-organized as

$$(d-\gamma s^*)\cdot E_H\left[f_{W;x_v}(d+H(s^*))\cdot H'(s^*)\right] + \gamma\cdot E_H\left[\overline{F}_{W;x_v}(d+H(s^*))\right] > 0,$$

implying that $\frac{\partial\,\psi(d,s^*;x_v)}{\partial s^*} > 0$. Hence, the denominator of (S.5), given by $\frac{\partial\,\psi(d,s;x_v)}{\partial s}|_{s=s_{\max}(d;x_v)}$, is always positive and the sign of $\frac{\partial\, s_{\max}(d;x_v)}{\partial d}$ is the opposite of the sign of the numerator in (S.5), given by $\frac{\partial\,\psi(d,s;x_v)}{\partial d}|_{s=s_{\max}(d;x_v)}$, where

$$\frac{\partial\,\psi(d,s;x_v)}{\partial d} = (d-\gamma s)\cdot E_H\left[f_{W;x_v}(d+H(s))\right] + E_H\left[F_{W;x_v}(d+H(s))\right] \tag{S.6}$$

(a) For any $d : B < d < d_{root}(0;x_v)$, we have $\frac{\partial\,\psi(d,s_{\max}(d;x_v);x_v)}{\partial d} > 0$ directly from (S.6). Hence, we conclude from (S.5) and the results above that $s_{\max}(d;x_v)$ is decreasing in d. Note that the equality, $s_{\max}(d_{root}(0;x_v);x_v) = 0$, simply follows by definition of $d_{root}(0;x_v)$.

(b) For any $d : 0 \le d < B$, we have $d < \gamma s$ (Property A.1(b)). Using (S.6), we conclude that $\frac{\partial\, s_{\max}(d;x_v)}{\partial d}$ can be positive or negative depending on its arguments. Hence, the internal solution $s_{\max}(d;x_v) \in (0,\overline{S})$ becomes nonincreasing in d if:

$$\frac{E_H\left[f_{W;x_v}(d+H(s_{\max}(d;x_v)))\right]}{E_H\left[F_{W;x_v}(d+H(s_{\max}(d;x_v)))\right]} \le \frac{1}{\gamma s_{\max}(d;x_v) - d},\ \forall d \in (0,B).$$

□

COROLLARY 2. *Consider **JM**. For any screening-centric strategy ($0 \le d < B$), if $\exists\, d' : 0 \le d < d' < B$ and $s_{\max}(d;x^{cr}_{min}) < s_{\max}(d';x^{cr}_{min})$, then solution $(d, s_{\max}(d;x^{cr}_{min}))$ cannot be in **JM** Pareto set.*

Proof of Corollary 2. Follows by Assumption 3 that $\frac{\partial \mathcal{G}(d,s;x_v)}{\partial s} > 0,\ \forall x_v \in \mathcal{S}(X_v), \forall d \geq 0, \forall s \geq 0.$ □

PROPOSITION 4. *For any incentive amount d, the optimal screening rate on* **JM** *Pareto set $(s^{*J}(d))$, and* **QM-2** *expected screening rate $(E_{X_v}[s^{*Q}(d;X_v)])$ respond to changes in population characteristics and institutional parameters as follows:*

(a) *As X_v increases in the usual stochastic order:*

(a-i) *Incentive-centric ($B < d \leq d_{root}(0;x_v)$): Both $s^{*J}(d)$ and $E_{X_v}[s^{*Q}(d;X_v)]$ decrease.*

(a-ii) *Screening-centric ($0 \leq d < B$): Both $s^{*J}(d)$ and $E_{X_v}[s^{*Q}(d;X_v)]$ increase.*

(b) *As γ increases, both $s^{*J}(d)$ and $E_{X_v}[s^{*Q}(d;X_v)]$ increase for any $d \neq B$.*

(c) *As B increases, both $s^{*J}(d)$ and $E_{X_v}[s^{*Q}(d;X_v)]$ increase for any $d \neq B$.*

(d) *As α increases, $s^{*J}(d)$ decreases while $E_{X_v}[s^{*Q}(d;X_v)]$ remains unchanged for any $d \neq B$.*

(e) *Consider $H(s) = \mathcal{K}s$. As the positive random variable $\mathcal{K}$ increases in the usual stochastic order:*

(e-i) *Incentive-centric ($B < d \leq d_{root}(0;x_v)$): Both $s^{*J}(d)$ and $E_{X_v}[s^{*Q}(d;X_v)]$ decrease.*

(e-ii) *Screening-centric ($0 \leq d < B$): Both $s^{*J}(d)$ and $E_{X_v}[s^{*Q}(d;X_v)]$ increase.*

Proof of Proposition 4. We first provide Lemma B.1 that will be used in the subsequent proofs.

LEMMA B.1. *For any $x_v \in \mathcal{S}(X_v)$, the roots, $s_{\max}(d;x_v)$ and $d_{root}(0;x_v)$, satisfy the following:*

(a) *As x_v increases:*

(a-i) *For any incentive-centric strategy ($B < d \leq d_{root}(0;x_v)$), $s_{\max}(d;x_v)$ decreases;*

(a-ii) *For any screening-centric strategy ($0 \leq d < B$), $s_{\max}(d;x_v)$ increases.*

(b) *As γ increases, $s_{\max}(d;x_v)$ is nonincreasing for all $0 \leq d < d_{root}(0;x_v)$.*

(c) *As B increases, $s_{\max}(d;x_v)$ increases for all $0 \leq d < d_{root}(0;x_v)$.*

(d) *As α increases:*

(d-i) *$s_{\max}(d;x_v)$ remains unchanged for all $0 \leq d \leq d_{root}(0;x_v)$;*

(d-ii) *For any incentive-centric strategy ($B < d \leq d_{root}(0;x_v)$), $s_{\max}(d;x^{cr}_{max})$ decreases, where $x^{cr}_{max} := F^{-1}_{X_v}(\alpha)$;*

(d-iii) *For any screening-centric strategy ($0 \leq d < B$), $s_{\max}(d;x^{cr}_{min})$ decreases, where $x^{cr}_{min} := \overline{F}^{-1}_{X_v}(\alpha)$.*

(e) *for $H(s) = \mathcal{K}s$, as the positive random variable $\mathcal{K}$ increases in the usual stochastic order:*

(e-i) *For any incentive-centric strategy ($B < d \leq d_{root}(0;x_v)$), $s_{\max}(d;x_v)$ decreases;*

(e-ii) *For any screening-centric strategy ($0 \leq d < B$), $s_{\max}(d;x_v)$ increases.*

Proof of Lemma B.1. First, recall that $\psi(d,s;x_v) = d \cdot E_H\left[F_{W;x_v}\big(d+H(s)\big)\right] + \gamma s \cdot E_H\left[\overline{F}_{W;x_v}\big(d+H(s)\big]\right) - B$. Using the implicit function theorem, we can write

$$\frac{\partial\, s_{\max}(d;x_v)}{\partial v} = -\frac{\frac{\partial\, \psi(d,s;x_v)}{\partial v}|_{s=s_{\max}(d;x_v)}}{\frac{\partial\, \psi(d,s;x_v)}{\partial s}|_{s=s_{\max}(d;x_v)}}, \tag{S.7}$$

for any internal solution $s_{\max}(d;x_v) \in (0,\overline{S})$, where v is any parameter of interest such as x_v, γ, or α. Recall from (S.1) in §A.3 and the proof of Proposition 3 that, given Assumption 4, we have

$$\frac{\partial\,\psi(d,s;x_v)}{\partial s} = (d-\gamma s)\cdot E_H\left[f_{W;x_v}(d+H(s))\cdot H'(s)\right] + \gamma\cdot E_H\left[\overline{F}_{W;x_v}\big(d+H(s)\big)\right] > 0. \tag{S.8}$$

Therefore, the denominator of (S.7) is always positive, and the sign of $\frac{\partial\, s_{\max}(d;x_v)}{\partial v}$ is the opposite of the sign of the numerator, given by $\frac{\partial\,\psi(d,s;x_v)}{\partial v}\big|_{s=s_{\max}(d;x_v)}$.

(a) As x_v increases, we know from Property A.2(a) that:

(a-i) $\frac{\partial}{\partial x_v}\psi(d,s;x_v)$ is positive for $d>B$, and therefore, $s_{\max}(d;x_v)$ decreases.

(a-ii) $\frac{\partial}{\partial x_v}\psi(d,s;x_v)$ is negative for $d<B$, and therefore, $s_{\max}(d;x_v)$ increases.

(b) We have

$$\frac{\partial}{\partial\gamma}\psi(d,s;x_v) = s\cdot E_H\left[\overline{F}_{W;x_v}\big(d+H(s)\big)\right] \geq 0, \tag{S.9}$$

which implies that $s_{\max}(d;x_v)$ is nonincreasing as γ increases.

(c) We have

$$\frac{\partial}{\partial B}\psi(d,s;x_v) = -1 < 0, \tag{S.10}$$

which implies that $s_{\max}(d;x_v)$ increases as B increases.

(d)

(d-i) We have $\frac{\partial\,\psi(d,s;x_v)}{\partial\alpha} = 0, \forall d$ since $\psi(d,s;x_v)$ is independent of α. Thus, as α increases, $s_{\max}(d;x_v)$ remains unchanged for all $0 \leq d \leq d_{root}(0;x_v)$.

(d-ii) For $d>B$, we know that $x^{cr}_{max} = F^{-1}_{X_v}(\alpha)$ and,

$$\frac{\partial\,\psi(d,s;x^{cr}_{max})}{\partial\alpha} = \frac{\partial\,\psi(d,s;x^{cr}_{max})}{\partial x^{cr}_{max}}\cdot\frac{\partial\,x^{cr}_{max}}{\partial\alpha}$$

$$= (d-\gamma s)\cdot\frac{\partial E_H\left[\,F_{W;x^{cr}_{max}}\big(d+H(s)\big)\right]}{\partial x^{cr}_{max}}\cdot\frac{\partial\,x^{cr}_{max}}{\partial\alpha}.$$

By definition, x^{cr}_{max} is increasing in α. Moreover, $E_H\left[\frac{\partial\,F_{W;x^{cr}_{max}}(d+H(s))}{\partial x^{cr}_{max}}\right] > 0$ by Assumption 2 and $d>B>\gamma s$ by Property A.1(b). Thus $\frac{\partial\,\psi(s_{\max}(d;x^{cr}_{max}))}{\partial\alpha} > 0$ and $s_{\max}(d;x^{cr}_{max})$ decreases in α.

(d-iii) For $d<B$, we know that $x^{cr}_{min} = \overline{F}^{-1}_{X_v}(\alpha)$ and,

$$\frac{\partial\,\psi(d,s;x^{cr}_{min})}{\partial\alpha} = (d-\gamma s)\cdot\frac{\partial\,E_H\left[F_{W;x^{cr}_{min}}\big(d+H(s)\big)\right]}{\partial x^{cr}_{min}}\cdot\frac{\partial\,x^{cr}_{min}}{\partial\alpha}.$$

By definition, x^{cr}_{min} is decreasing in α. Moreover, $E_H\left[\frac{\partial\,F_{W;x^{cr}_{min}}(d+H(s))}{\partial x^{cr}_{min}}\right] > 0$ by Assumption 2 and $d<B<\gamma s$ by Property A.1(b). Thus $\frac{\partial\,\psi(s_{\max}(d;x^{cr}_{min}))}{\partial\alpha} > 0$ and $s_{\max}(d;x^{cr}_{min})$ decreases in α.

(e) Consider $\mathcal{K}' <_{st} \mathcal{K}$ with corresponding $s_{\max}(d,\mathcal{K}';x_v)$, $\psi(d,s,\mathcal{K}';x_v)$, $s_{\max}(d,\mathcal{K};x_v)$, and $\psi(d,s,\mathcal{K};x_v)$. Letting $\xi(s,k) := (d-\gamma s)\cdot F_{W;x_v}\big(d+ks\big) - (B-\gamma s)$, we have

$$\frac{\partial \xi(s,k)}{\partial k} = (d-\gamma s)\cdot f_{W;x_v}\big(d+ks\big)\cdot k. \tag{S.11}$$

(e-i) For $d > B$, we have from (S.11) that $\xi(s,k)$ increases in k. Therefore, by definitions of $s_{\max}(\cdot)$ and $\psi(d,s;x_v)$ and from Remark A.1, we have

$$\begin{aligned} 0 = \psi(d, s_{\max}(d,\mathcal{K};x_v),\mathcal{K};x_v) &= E_{\mathcal{K}}\left[\xi(s_{\max}(d,\mathcal{K};x_v),\mathcal{K})\right] \\ &> E'_{\mathcal{K}}\left[\xi(s_{\max}(d,\mathcal{K};x_v),\mathcal{K}')\right] \\ &= \psi(d, s_{\max}(d,\mathcal{K};x_v);\mathcal{K}',x_v). \end{aligned}$$

Recall from Property A.2 (b-ii) that $\lim_{s\to\infty}\psi(d,s,\mathcal{K};x_v)\to\infty$, which together with $\psi(d, s_{\max}(d,\mathcal{K};x_v);\mathcal{K}',x_v) < 0$ implies, by intermediate value theorem, that there exists an $s_{root}(d,\mathcal{K}';x_v) > s_{\max}(d,\mathcal{K};x_v)$, which gives $s_{\max}(d,\mathcal{K}';x_v) > s_{\max}(d,\mathcal{K};x_v)$.

(e-ii) For $d > B$, we have from (S.11) that $\xi(s,k)$ decreases in k. Following similar steps to the proof of part (e-ii), by definitions of $s_{\max}(\cdot)$ and $\psi(d,s;x_v)$ and from Remark A.1, we have

$$\begin{aligned} 0 = \psi(d, s_{\max}(d,\mathcal{K}';x_v),\mathcal{K}';x_v) &= E'_{\mathcal{K}}\left[\xi(s_{\max}(d,\mathcal{K}';x_v),\mathcal{K}')\right] \\ &> E_{\mathcal{K}}\left[\xi(s_{\max}(d,\mathcal{K}';x_v),\mathcal{K})\right] \\ &= \psi(d, s_{\max}(d,\mathcal{K}';x_v);\mathcal{K},x_v). \end{aligned}$$

Therefore, by Property A.2 (b-ii) and intermediate value theorem, there exists an $s_{root}(d,\mathcal{K};x_v) > s_{\max}(d,\mathcal{K}';x_v)$, which gives $s_{\max}(d,\mathcal{K};x_v) > s_{\max}(d,\mathcal{K}';x_v)$.

□

Next, we prove Proposition 4.

(a) As X_v increases in stochastic order:

(a-i) For $d > B$, by Proposition 1, the constraint is tight at x^{cr}_{max} and thus, using Assumption 4, $s^{*J}(d) = s_{\max}(d;x^{cr}_{max})$. In addition, as X_v increases in stochastic order, $F^{-1}_{X_v}(\alpha) = x^{cr}_{max}$ increases (for a given α) (Definition A.1, §A.2). Observe that $F^{-1}_{X_v}(\alpha)$ is what changes in $\Psi(d,s,X_v)$ as X_v increases in stochastic order or as α increases; in both cases $F^{-1}_{X_v}(\alpha)$ increases. Thus, the direction of the behavior of $s_{\max}(d;x^{cr}_{max})$ as X_v increases in stochastic order must be the same as the direction as α increases. Since Lemma B.1(d-ii) establishes that $s_{\max}(d;x^{cr}_{max})$ decreases in α for $d > B$, $s^{*J}(d)$ decreases as X_v increases in usual stochastic order. Finally, using Lemma B.1(a-i), which establishes that $s_{\max}(d;x_v)$ decreases in x_v for $d > B$, and Remark A.1(b), we deduce that $E[s^{*Q}(d;X_v)] = E_{X_v}[s_{\max}(d;x_v)]$ decreases as X_v increases in usual stochastic order.

(a-ii) For $d < B$, by Proposition 1, the constraint is tight at x^{cr}_{min} and thus, using Assumption 4, $s^{*J} = s_{\max}(d; x^{cr}_{min})$. In addition, as X_v increases in stochastic order, $F^{-1}_{X_v}(1-\alpha) = x^{cr}_{min}$ increases (for a given α) (Definition A.1, §A.2). Because $F^{-1}_{X_v}(1-\alpha)$ is what changes in $\Psi(d, s, X_v)$ as X_v increases in stochastic order and as α increases, we deduce that the behavior of $s_{\max}(d; x^{cr}_{min})$ should be the opposite of the one we get as α increases for $d < B$. Hence, knowing that Lemma B.1(d-iii) states that $s_{\max}(d; x^{cr}_{min})$ decreases in α for $d < B$, we conclude that $s^{*J}(d)$ increases as X_v increases in usual stochastic order. Finally, similar to the proof of part (e-i), using Lemma B.1(a-ii), which establishes that $s_{\max}(d; x_v)$ increases in x_v for $d < B$, and Remark A.1(b), $E[s^{*Q}(d; X_v)] = E_{X_v}[s_{\max}(d; x_v)] = \int_{x_v \in \mathcal{S}(X_v)} s_{\max}(d; x_v) \cdot f_{X_v}(x_v)\, dx_v$ increases as X_v increases in usual stochastic order.

(b) By Proposition 1, **JM**'s constraint is tight when $d \neq B$. By Assumption 4, $s^{*J}(d) = s_{\max}(d; x^{cr}_{max})$ for $d > B$ and $s^{*J}(d) = s_{\max}(d; x^{cr}_{min})$ for $d < B$, where x^{cr}_{min} and x^{cr}_{max} are independent of γ. Further, by Assumption 4 and the definition of expected value, $E_{X_v}[s^{*Q}(d; X_v)] = E_{X_v}[s_{\max}(d; x_v)], \forall d \neq B$. Thus, the result follows from Lemma B.1(b).

(c) By Proposition 1, **JM**'s constraint is tight when $d \neq B$. By Assumption 4, $s^{*J}(d) = s_{\max}(d; x^{cr}_{max})$ for $d > B$ and $s^{*J}(d) = s_{\max}(d; x^{cr}_{min})$ for $d < B$, where x^{cr}_{min} and x^{cr}_{max} are independent of B. Further, by Assumption 4 and the definition of expected value, ${}_{X_v}E[s^{*Q}(d; X_v)] = E_{X_v}[s_{\max}(d; x_v)], \forall d \neq B$. Thus, the result follows from Lemma B.1(c).

(d) By Proposition 1, **JM**'s constraint is tight when $d \neq B$. By Assumption 4, $s^{*J}(d) = s_{\max}(d; x^{cr}_{max})$ for $d > B$ and $s^{*J}(d) = s_{\max}(d; x^{cr}_{min})$ for $d < B$. Further, by Assumption 4 and the definition of expected value, $E_{X_v}[s^{*Q}(d; X_v)] = E_{X_v}[s_{\max}(d; x_v)], \forall d \neq B$. Thus, the result follows from Lemma B.1(d).

(e) By Proposition 1, **JM**'s constraint is tight when $d \neq B$. By Assumption 4, $s^{*J}(d) = s_{\max}(d; x^{cr}_{max})$ for $d > B$ and $s^{*J}(d) = s_{\max}(d; x^{cr}_{min})$ for $d < B$, where x^{cr}_{min} and x^{cr}_{max} are independent of $H(s)$. Further, by Assumption 4 and the definition of expected value, $E_{X_v}[s^{*Q}(d; X_v)] = E_{X_v}[s_{\max}(d; x_v)], \forall d \neq B$. Thus, the result follows from Lemma B.1(e).

□

PROPOSITION 5. *Consider the zero-h special case. For any $x_v \in \mathcal{S}(X_v)$ and $0 \leq d \leq d_{root}(0; x_v)$, $s_{\max}(d; x_v)^{(h=0)} = \frac{B - d \cdot F_{W;x_v}(d)}{\gamma \cdot \overline{F}_{W;x_v}(d)}$ is the unique solution to, $\psi^{(h=0)}(d, s; x_v) = 0$. Then:*

(a) *For any incentive-centric strategy ($B < d \leq d_{root}(0; x_v)$), $s_{\max}(d; x_v) < s_{\max}(d; x_v)^{(h=0)}$.*

(b) *For any screening-centric strategy ($0 \leq d < B$), $s_{\max}(d; x_v) > s_{\max}(d; x_v)^{(h=0)}$.*

Proof of Proposition 5. When $h(s) = 0$, the solution $s_{\max}(d; x_v)^{(h=0)}$ to $\psi(d, s; x_v) = 0$, can be explicitly derived by rewriting $\psi(d, s; x_v) = 0$ as

$$d \cdot F_{W;x_v}\big(d\big) + \gamma s \cdot \overline{F}_{W;x_v}\big(d\big) - B = 0. \tag{S.12}$$

Reorganizing (S.12) leads to the following explicit characterization:

$$s_{\max}(d;x_v)^{(h=0)} = \frac{B - d \cdot F_{W;x_v}\big(d\big)}{\gamma \cdot \overline{F}_{W;x_v}\big(d\big)}, \tag{S.13}$$

which also implies that for given d and x_v, $s_{\max}(d;x_v)^{(h=0)}$ is unique.

Moreover, the following holds for the *general* and *zero-h* cases (see (S.1) in §A.3):

$$\psi(d, s_{\max}(d;x_v);x_v) = 0,$$
$$\implies (d - \gamma s_{\max}(d;x_v)) \cdot F_{W;x_v}\big(d + H(s_{\max}(d;x_v))\big) - (B - \gamma s_{\max}(d;x_v)) = 0, \tag{S.14a}$$

$$\psi(d, s_{\max}(d;x_v)^{(h=0)};x_v)^{(h=0)} = 0,$$
$$\implies (d - \gamma s_{\max}(d;x_v)^{(h=0)}) \cdot F_{W;x_v}\big(d\big) - (B - \gamma s_{\max}(d;x_v)^{(h=0)}) = 0, \tag{S.14b}$$

$$\frac{\partial}{\partial s}\psi(d,s;x_v) = (d - \gamma s) \cdot f_{W;x_v}(d + H(s)) \cdot H'(s) + \gamma \cdot \overline{F}_{W;x_v}\big(d + H(s)\big), \tag{S.14c}$$

$$\frac{\partial}{\partial s}\psi(d,s;x_v)^{(h=0)} = \gamma \cdot \overline{F}_{W;x_v}\big(d\big). \tag{S.14d}$$

Therefore,

(a) For any $B < d \leq d_{root}(0;x_v)$, we have $\gamma s < B < d$ (Property A.1(b)) and $\psi(d,0;x_v) = \psi(d,0;x_v)^{(h=0)} = d \cdot F_{W;x_v}\big(d\big) - B \leq 0$. Further, from (S.14c) and (S.14d), we have $\frac{\partial}{\partial s}\psi(d,s;x_v) > \frac{\partial}{\partial s}\psi(d,s;x_v)^{(h=0)} > 0, \forall s > 0$. Therefore, since $\psi(d,0;x_v) = \psi(d,0;x_v)^{(h=0)} \leq 0$, and both $\psi(d,s;x_v)$ and $\psi(d,s;x_v)^{(h=0)}$ are increasing in s such that $\frac{\partial}{\partial s}\psi(d,s;x_v) > \frac{\partial}{\partial s}\psi(d,s;x_v)^{(h=0)}$, (S.14a) and (S.14b) imply that $s_{\max}(d;x_v) < s_{\max}(d;x_v)^{(h=0)}$.

(b) For any $d < B$, we have $d < B < \gamma s$ Property A.1(b), and $\psi(d,0;x_v) = \psi(d,0;x_v)^{(h=0)} = d \cdot F_{W;x_v}\big(d\big) - B \leq 0$. Further, from (S.14c) and (S.14d), we have $\frac{\partial}{\partial s}\psi(d,s;x_v)^{(h=0)} > 0$ and $\frac{\partial}{\partial s}\psi(d,s;x_v)^{(h=0)} > \frac{\partial}{\partial s}\psi(d,s;x_v)$. Therefore, since $\psi(d,0;x_v) = \psi(d,0;x_v)^{(h=0)} \leq 0$, and $\psi(d,s;x_v)^{(h=0)}$ is increasing in s whereas $\psi(d,s;x_v)$ may increase or decrease such that $\frac{\partial}{\partial s}\psi(d,s;x_v)^{(h=0)} > \frac{\partial}{\partial s}\psi(d,s;x_v)$, (S.14a) and (S.14b) imply that $s_{\max}(d;x_v) > s_{\max}(d;x_v)^{(h=0)}$.

□

Appendix C: Additional Results

C.1. Additional Results on $s_{root}(\cdot)$ Function

PROPOSITION C.1. *For any $0 \leq d \leq d_{root}(0;x_v)$ and $x_v \in \mathcal{S}(X_v)$, if:*

$$\frac{E_H\left[\overline{F}_{W;x_v}\big(d+H(s)\big]\cdot E_H\left[\frac{\partial f_{W;x_v}(d+H(s))}{\partial s}\cdot H'(s)^2 + f_{W;x_v}(d+H(s))\cdot H''(s)\right]\right.}{E_H\left[f_{W;x_v}(d+H(s))\cdot H'(s)\right]^2} \geq -2,\ \forall s \in [0,\infty), \tag{S.15}$$

then $\psi(d,s;x_v)$ is unimodal in s. In this case:

(a) *$s_{root}(d;x_v)$ has at most two solutions for $d : 0 \leq d < B$.*

(b) *$s_{root}(d;x_v)$ is unique for $d : B \leq d \leq d_{root}(0;x_v)$.*

Proof of Proposition C.1. First, recall that $s_{root}(d;x_v) := \{s \geq 0 :\ \psi(d,s;x_v) = 0\}$. Let $s_{FOC} := \{s \geq 0 :\ \psi'(s) = 0\}$, i.e., the set of *First-order Condition (FOC)* solutions.

From (S.2a) in §A.3, we deduce the following:

$$\begin{aligned} \forall s \in s_{FOC} :\ \psi'(s) = 0 &\Leftrightarrow s = d/\gamma + \frac{E_H\left[\overline{F}_{W;x_v}\big(d+H(s)\big)\right]}{E_H\left[f_{W;x_v}(d+H(s))\cdot H'(s)\right]} \\ &\Leftrightarrow s - \frac{E_H\left[\overline{F}_{W;x_v}\big(d+H(s)\big)\right]}{E_H\left[f_{W;x_v}(d+H(s))\cdot H'(s)\right]} = d/\gamma. \end{aligned} \tag{S.16}$$

As a result, we can find the s_{FOC} by studying (S.16) and its roots. The derivative of the LHS of (S.16) in s is positive if and only if:

$$\frac{E_H\left[\overline{F}_{W;x_v}\big(d+H(s)\big]\cdot E_H\left[\frac{\partial f_{W;x_v}(d+H(s))}{\partial s}\cdot H'(s)^2 + f_{W;x_v}(d+H(s))\cdot H''(s)\right]\right.}{E_H\left[f_{W;x_v}(d+H(s))\cdot H'(s)\right]^2} \geq -2$$

Therefore, when inequality (S.15) holds $\forall s \in [0,\infty)$, then LHS is monotone increasing in s_{FOC}, whereas the RHS is a constant. Hence, the FOC solution is unique, i.e., $|s_{FOC}| = 1$ and $\psi(d,s;x_v)$ is unimodal in s.

(a) Because $\psi(d,s;x_v)$ is unimodal in s, it has at most two s-roots, i.e., $|s_{root}(d;x_v)| \in \{0,1,2\}$ and $s_{root}(d;x_v)$ has at most two solutions for a given d, including $d : 0 \leq d < B$.

(b) If $d = B$, then $s_{root}(d;x_v)$ is unique and is equal to $\frac{B}{\gamma}$ (see Property A.1(b)).
For $d : B < d \leq d_{root}(0;x_v)$, we have that $\psi(d,0;x_v) := d\cdot F_{W;x_v}\big(d\big) - B \leq 0$, and $\psi(d,d/\gamma;x_v) := (d-\gamma s)\cdot E_H\left[F_{W;x_v}\big(d+H(s)\big)\right] - (B-\gamma s) = d - B > 0$. Because $\psi(d,s;x_v)$ is increasing in $s \in [0,\frac{d}{\gamma}]$ (Property A.2(b-i)), we conclude by the intermediate value theorem that one root exists in this region. Moreover, we have $\lim_{s\to\infty}\psi(d,s;x_v) \to \infty$ (Property A.2(b-ii)). Therefore, when equation (S.15) is satisfied, $|s_{root}(d;x_v)|$ can only be equal to 1 since $\psi(d,s;x_v)$ is unimodal (see above). Equivalently, since $\lim_{s\to\infty}\psi(d,s;x_v) \to \infty$, we deduce that $\psi(d,s;x_v) > 0, \forall s > \frac{d}{\gamma}$, and therefore, no other root exists in this region.

□

C.2. Value of Information

To quantify the value of information on vaccination coverage in the sequential model **QM-2**, we compare, for each incentive amount (d), screening rates and health rewards in the Pareto sets of the joint and sequential models in Proposition C.2. Let $s^{*Q}(d,X_v^N;X_v)$ denote the optimal screening rate under **QM-2**; and $\mathcal{G}^{*Q}(d,X_v^N;X_v) = \mathcal{G}\left(d,s^{*Q}(d,X_v^N;X_v),X_v^N\right)$ and $\mathcal{G}^{*J}(d,X_v^N;X_v) = \mathcal{G}\left(d,s^{*J}(d),X_v^N;X_v\right)$ resp. denote the optimal health rewards under **QM-2** and **JM** when the Nature's initial vaccination coverage is X_v^N but the planner models it as X_v.

PROPOSITION C.2. *For any incentive- ($B < d \leq d_{root}(0;x_v)$) and screening-centric ($0 \leq d < B$) strategy in the Pareto sets, optimal screening rates ($s^{*Q}(\cdot)$ and $s^{*J}(\cdot)$), health rewards ($\mathcal{G}^{*Q}(\cdot)$ and $\mathcal{G}^{*J}(\cdot)$), and budget-compliance probabilities ($Pr\big\{\Psi(\cdot)\leq 0\big\}$) for **JM** and **QM-2** compare as follows:*

(a) $Pr\{s^{*Q}(d, X_v^N; X_V) \geq s^{*J}(d)\} \in \Big[\min\{F_{X_v^N}(x_{max}^{cr}), \overline{F}_{X_v}(x_{min}^{cr})\}, \max\{F_{X_v^N}(x_{max}^{cr}), \overline{F}_{X_v}(x_{min}^{cr})\Big]\}$, *and further,* $Pr\{s^{*Q}(d, X_v^N) \geq s^{*J}(d)\} = \alpha$ *for* $X_v^N = X_v$.

(b) $Pr\{\mathcal{G}^{*Q}(d, X_v^N; X_V) \geq \mathcal{G}^{*J}(d, X_v^N; X_V)\} \in \Big[\min\{F_{X_v^N}(x_{max}^{cr}), \overline{F}_{X_v}(x_{min}^{cr})\}], \max\{F_{X_v^N}(x_{max}^{cr}), \overline{F}_{X_v}(x_{min}^{cr})\Big]\}$, *and, further* $Pr\{\mathcal{G}^{*Q}(d, X_v^N) \geq \mathcal{G}^{*J}(d, X_v^N)\} = \alpha$ *for* $X_v^N = X_v$.

(b-i) *Further,* $Pr\{\Psi(d, s^{*Q}(d, X_v), X_v) \leq 0\} = 1$ *for any screening-centric strategy* ($d < B$) *in* ***QM-2***.

(c) *The difference,* $E_{X_v}[s^{*Q}(d, X_v)] - s^{*J}(d)$, *is increasing in* α.

Proof of Proposition C.2. We first provide Lemmas C.2 and C.3 that will be used in the subsequent proofs.

LEMMA C.2. *For any incentive-centric strategy* ($B < d \leq d_{root}(0; x_v)$)*, the following relationships hold:*

$$
\begin{aligned}
&s^{*Q}(d; x_v) > s^{*J}(d) > 0, && \text{if } 0 < x_v < x_{max}^{cr}\\
&s^{*Q}(d; x_v) = s^{*J}(d) > 0, && \text{if } x_v = x_{max}^{cr}\\
&s^{*J}(d) > s^{*Q}(d; x_v) > 0, && \text{if } x_{max}^{cr} < x_v < x_{max}(d, 0)\\
&s^{*J}(d) > s^{*Q}(d; x_v) = 0, && \text{if } x_v \geq x_{max}(d, 0).
\end{aligned}
$$

Proof of Lemma C.2. By Lemma B.1(a-i), as x_v increases, $s_{\max}(d; x_v)$ decreases for $d > B$. In addition, Corollary 1 establishes that $\forall(d, s_{\max}(d; x_{max}^{cr})) : B < d \leq d_{root}(0; x_{max}^{cr})$, the **JM** Pareto set has $s_{\max}(d; x_{max}^{cr})$, which exists given Assumption 4. Thus, while x_{max}^{cr} is used to get the screening solution for **JM**, the actual x_v realization is used to get the solution in **QM**-2. Hence, $s^{*Q}(d; x_v) = s_{\max}(d; x_v)$ and $s^{*J}(d) = s_{\max}(d; x_{max}^{cr})$. As a result,

- For $0 < x_v < x_{max}^{cr}$, using Lemma B.1(a-i) and that $s^{*J}(d) = s_{\max}(d; x_{max}^{cr})$, we have $s^{*Q}(d; x_v) > s^{*J}(d) > 0$.
- For $x_v = x_{max}^{cr}$, using that $s^{*J}(d) = s_{\max}(d; x_{max}^{cr}) = s_{\max}(d; x_v) = s^{*Q}(d; x_v)$, we have $s^{*Q}(d; x_v) = s^{*J}(d)$. In addition, using Proposition 2(b) and that $d < d_{root}(0; x_{max}^{cr})$, we have $s^{*Q}(d; x_v) > 0$.
- For $x_{max}^{cr} < x_v < x_{max}(d, 0)$, using Lemma B.1(a-i) and that $s^{*J}(d) = s_{\max}(d; x_{max}^{cr})$, we have $s^{*J}(d) > s^{*Q}(d; x_v)$. In addition, using Proposition 2(b) and that $x_v < x_{max}(d, 0)$, or equivalently, $d < d_{root}(0; x_{max}^{cr})$, we have $s^{*Q}(d; x_v) > 0$.
- For $x_v \geq x_{max}(d, 0)$, using Lemma B.1(a-i) and that $s^{*J}(d) = s_{\max}(d; x_{max}^{cr})$, we have $s^{*J}(d) > s^{*Q}(d; x_v)$. In addition, using Proposition 2(a) and that $x_v \geq x_{max}(d, 0)$, we have $s^{*Q}(d; x_v) = 0$.

□

LEMMA C.3. *For any screening-centric strategy* ($0 \leq d < B$)*, we have:*

$$
\begin{aligned}
&s^{*Q}(d; x_v) < s^{*J}(d), && \text{if } 0 < x_v < x_{min}^{cr}\\
&s^{*Q}(d; x_v) = s^{*J}(d), && \text{if } x_v = x_{min}^{cr}\\
&s^{*Q}(d; x_v) > s^{*J}(d), && \text{if } x_v > x_{min}^{cr}.
\end{aligned}
$$

Proof of Lemma C.3. The proof is similar to that of Lemma C.2. Lemma B.1(a-ii) establishes that as x_v increases, $s_{\max}(d;x_v)$ increases for $d<B$. In addition, by Corollary 1, $\forall(d, s_{\max}(d;x^{cr}_{min})) : 0 \le d < B$, the **JM** Pareto set has $s_{\max}(d;x^{cr}_{min})$, which exists given Assumption 4. Thus, the screening solution for **JM** is obtained at x^{cr}_{min}, whereas the actual x_v realization is used to get the solution in **QM**-2. Hence, $s^{*Q}(d;x_v) = s_{\max}(d;x_v) > \frac{B}{\gamma}$ and $s^{*J}(d) = s_{\max}(d;x^{cr}_{min}) > \frac{B}{\gamma}$. As a result,

- For $0 < x_v < x^{cr}_{min}$, using Lemma B.1(a-ii) and that $s^{*J}(d) = s_{\max}(d;x^{cr}_{min})$, $s^{*Q}(d;x_v) < s^{*J}(d)$.
- For $x_v = x^{cr}_{min}$, using that $s^{*J}(d) = s_{\max}(d;x^{cr}_{min}) = s_{\max}(d;x_v) = s^{*Q}(d;x_v)$, $s^{*Q}(d;x_v) = s^{*J}(d)$.
- For $x_v > x^{cr}_{min}$, using Lemma B.1(a-ii) and that $s^{*J}(d) = s_{\max}(d;x^{cr}_{min})$, $s^{*Q}(d;x_v) > s^{*J}(d)$.

□

Next, we prove Proposition C.2.

(a) We know from Lemma C.2 and Lemma C.3 that $s^{*Q}(d;x_v) \ge s^{*J}(d)$ for $0 < x_v \le x^{cr}_{max}$ (when $d > B$) and $x_v \ge x^{cr}_{min}$ (when $d < B$), resp.. Next, using the law of total probability, we get

$$\begin{aligned} Pr\big\{s^{*Q}(d,X^N_v) \ge s^{*J}(d)\big\} &= Pr\big\{0 < X^N_v \le x^{cr}_{max}; d > B\big\} \cdot Pr\big\{d > B\big\} \\ &\quad + Pr\big\{X^N_v \ge x^{cr}_{min}; d < B\big\} \cdot Pr\big\{d < B\big\} \\ &= Pr\big\{0 < X^N_v \le x^{cr}_{max}\big\} \cdot Pr\big\{d > B\big\} \\ &\quad + Pr\big\{X^N_v \ge x^{cr}_{min}\big\} \cdot (1 - Pr\big\{d > B\big\}) \end{aligned}$$

Thus, we have

$$Pr\big\{s^{*Q}(d,X^N_v) \ge s^{*J}(d)\big\} \in \Big[\min\{F_{X^N_v}(x^{cr}_{max}), \overline{F}_{X_v}(x^{cr}_{min}) \ , \max\{F_{X^N_v}(x^{cr}_{max}), \overline{F}_{X_v}(x^{cr}_{min})\Big].$$

Further, by definition of x^{cr}_{max} and x^{cr}_{min}, we have $Pr\big\{0 < X^N_v \le x^{cr}_{max}; d > B\big\} = Pr\big\{X^N_v \ge x^{cr}_{min}; d < B\big\} = \alpha$ for $X^N_v = X_v$, which implies that $Pr\big\{s^{*Q}(d,X^N_v) \ge s^{*J}(d)\big\} = \alpha$ for $X^N_v = X_v$.

(b) Because $Pr\big\{s^{*Q}(d,X^N_v) \ge s^{*J}(d)\big\} \in \Big[\min\{F_{X^N_v}(x^{cr}_{max}), \overline{F}_{X_v}(x^{cr}_{min}) \ , \max\{F_{X^N_v}(x^{cr}_{max}), \overline{F}_{X_v}(x^{cr}_{min})\Big]$ (Proposition C.2(a)) and $\mathcal{G}(.)$ is increasing in s for all $s \ge 0$ (Assumption 3), we have that $Pr\big\{\mathcal{G}^{*Q}(d,X_v) \ge \mathcal{G}^{*J}(d,X_v)\big\} \in \Big[\min\{F_{X^N_v}(x^{cr}_{max}), \overline{F}_{X_v}(x^{cr}_{min}) \ , \max\{F_{X^N_v}(x^{cr}_{max}), \overline{F}_{X_v}(x^{cr}_{min})\Big]$. Further, for $X^N_v = X_v$, we have $Pr\big\{s^{*Q}(d,X^N_v) \ge s^{*J}(d)\big\} = \alpha$ (Proposition C.2(a)), which implies that $Pr\big\{\mathcal{G}^{*Q}(d,X_v) \ge \mathcal{G}^{*J}(d,X_v)\big\} = \alpha$ for $X^N_v = X_v$.

(c) For any given d, we know from Lemma B.1(d) that $s^{*J}(d)$ decreases in α while $s^{*Q}(d;x_v)$ (and hence $E_{X_v}[s^{*Q}(d,X_v)]$) remains unchanged in α. As a result, the difference, $E_{X_v}[s^{*Q}(d,X_v)] - s^{*J}(d)$ is increasing in α.

□

Importantly, the value of information, in terms of attaining a higher health reward, does not necessarily increase as X_v becomes more variable (while keeping its mean constant), see Remark C.1. We can show, however, that for a screening (incentive)-centric strategy, if X_v becomes more variable through a heavier right tail after x^{cr}_{min} (heavier left tail prior to x^{cr}_{max}), the value of information increases (Corollary C.1).

REMARK C.1. The behavior of $s_{\max}(d; x_{max}), \forall B < d < d_{root}(0; x^{cr}_{max})$, and $s_{\max}(d; x_{min}), \forall d < B$, are not necessarily monotone in the second-order stochastic dominance (increasing concave order (icv)) of X_v.

Counter-example supporting Remark C.1. Remark C.1 follows because $X_v \geq_{icv} X_v\prime$ does not imply an ordering of their CDFs, which can potentially cross over multiple times. Consider $F_{X_v}(x) = 10x^3 - 15x^4 + 6x^5$ and $F_{X'_v}(x) = 3x^2 - 2x^3$, for $x \geq 0$. We have that $X_v \geq_{icv} X'_v$, as $\int_0^x [F_{X'_v}(t) - F_{X_v}(t)]\, dt \geq 0$. By construction, $\exists \tilde{x} \in \mathcal{S}(X_v)$ such that $F_{X_v}(\tilde{x}) = F_{X'_v}(\tilde{x})$. Moreover, $F_{X_v}(x) < F_{X'_v}(x), \forall x < \tilde{x}$, and $F_{X'_v}(x) < F_{X_v}(x), \forall x > \tilde{x}$. Consequently, it is possible to have $x_{max} = F^{-1}_{X_v}(\alpha) \geq x'_{max} = F^{-1}_{X'_v}(\alpha)$, and vice versa. The same argument extends to $x_{min} = \overline{F}^{-1}_{X_v}(\alpha)$ versus $x'_{min} = \overline{F}^{-1}_{X'_v}(\alpha)$. Then, Remark C.1 follows because $s_{\max}(d; x_v)$ is monotone in x_v (Lemma B.1(a)). □

COROLLARY C.1. (a) *Under a screening-centric strategy ($d < B$), for any X_V and X'_V such that $X_V \overset{d}{=} X'_V + Z$ for some non-negative random variable Z having $E[Z; X'_V = x'_v] = 0$, for any $x'_v \leq x^{cr}_{min} := \overline{F}^{-1}_{X_v}(\alpha)$, where $\overset{d}{=}$ means "has the same distribution as" (alt. "equal in distribution"), we have $E[s^{*Q}(d, X_v) - s^{*J}(d)] \geq E[s^{*Q}(d, X'_v) - s^{*J}(d)]$ and $E[\mathcal{G}(d, s^{*Q}(d, X_v)) - \mathcal{G}(d, s^{*J}(d))] \geq E[\mathcal{G}(d, s^{*Q}(d, X'_v)) - \mathcal{G}(d, s^{*J}(d))]$.*

(b) *Under an incentive-centric strategy ($d > B$), for any X_V and X'_V such that $X_V \overset{d}{=} X'_V - Z$ for some non-negative random variable Z having $E[Z; X'_V = x'_v] = 0$, for any $x'_v \geq x^{cr}_{max} := F^{-1}_{X_v}(\alpha)$, we have $E[s^{*Q}(d, X_v) - s^{*J}(d)] \geq E[s^{*Q}(d, X'_v) - s^{*J}(d)]$ and $E[\mathcal{G}(d, s^{*Q}(d, X_v)) - \mathcal{G}(d, s^{*J}(d))] \geq E[\mathcal{G}(d, s^{*Q}(d, X'_v)) - \mathcal{G}(d, s^{*J}(d))]$.*

Proof of Corollary C.1.

(a) By construction, X_v and X'_v have the same distribution from 0 to x^{cr}_{min} and then the distribution of X_v is spread to the right. As a result, $s^{*J}(d) = s_{\max}(d; x^{cr}_{min})$ remains the same under X_v and X'_v. However, by Proposition 4(e-ii), $E[s^{*Q}(d; X_v)] = E_{X_v}[s_{\max}(d; x_v)]$ increases as X_v increases in stochastic order (which is comparable to what happens for all $x_v > x^{cr}_{min}$), given Assumption 4. Hence, $E[s^{*Q}(d, X_v) - s^{*J}(d)] \geq E[s^{*Q}(d, X'_v) - s^{*J}(d)]$. Then $E[\mathcal{G}(d, s^{*Q}(d, X_v)) - \mathcal{G}(d, s^{*J}(d))] \geq E[\mathcal{G}(d, s^{*Q}(d, X'_v)) - \mathcal{G}(d, s^{*J}(d))]$ follows by Assumption 3.

(b) By construction, X_v and X'_v have the same distribution from x^{cr}_{max} to 1 and then the distribution of X_v is spread to the left. As a result, $s^{*J}(d) = s_{\max}(d; x^{cr}_{max})$ remains the same under X_v and X'_v. However, by Proposition 4(e-i), $E[s^{*Q}(d; X_v)] = E_{X_v}[s_{\max}(d; x_v)]$ increases as X_v decreases in stochastic order (which is comparable to what happens for all $x_v < x^{cr}_{max}$), given Assumption 4. Hence, $E[s^{*Q}(d, X_v) - s^{*J}(d)] \geq E[s^{*Q}(d, X'_v) - s^{*J}(d)]$. Then, $E[\mathcal{G}(d, s^{*Q}(d, X_v)) - \mathcal{G}(d, s^{*J}(d))] \geq E[\mathcal{G}(d, s^{*Q}(d, X'_v)) - \mathcal{G}(d, s^{*J}(d))]$ follows by Assumption 3.

□

Appendix D: Details of the Logistic Growth Model (§5)

We first provide the notation used in this section.

- P: Campus population
- K: Carrying capacity, i.e., number of susceptible individuals
- I_T: Cumulative number of infections at time T (the end of the semester)
- I_0 ($I_0 < K$): Initial number of infections
- r: infection growth rate
- R_0: Basic reproduction number, i.e., the expected number of secondary infections produced by a single infected individual in a fully susceptible population
- R_e: Effective reproduction number, i.e., the expected number of secondary infections produced by a single infected individual in a population at any specific time, influenced by the community's existing level of immunity
- β: Vaccine efficacy, i.e., the conditional probability that a vaccinated individual will not get infected nor infectious, given exposure
- τ: Generation time of the infection, i.e., average time between infection of a primary case and infection of its secondary cases (e.g., if the transmitter shows symptoms on Day 1 and the transmitted on Day 5, then $\tau = 4$).
- t_I: Infectious period, i.e., the time an individual remains infectious after exposure
- $p(s)$: Isolation intensity, i.e., expected proportion of the infectious period during which an unvaccinated, infected individual is in isolation, hence cannot transmit the infection, due to routine screening at rate s

D.1. Proof for the LGM Model (§5.1)

The underlying assumptions for the logistic growth model (LGM) are described in §5.1. Under LGM, the cumulative number of infections by time T, I_T, follows (Tsoularis and Wallace 2002):

$$I_T = \frac{K}{1 + \left(\frac{K - I_0}{I_0}\right) e^{-rT}}. \tag{S.17}$$

Infection growth rate r and effective reproduction number R_e are related (Wallinga and Lipsitch 2007):

$$r = \frac{\ln(R_e)}{\tau}. \tag{S.18}$$

LEMMA D.1. *The function $p(s)$ is a nondecreasing function of screening rate s:*

$$p(s) = \begin{cases} 0, & \text{if } s = 0 \\ \frac{t_I \cdot s}{2T}, & \text{if } 1 \le s \le \frac{T}{t_I} \\ 1 - \frac{T}{2t_I \cdot s}, & \text{if } \frac{T}{t_I} \le s \le T \end{cases}.$$

Proof of Lemma D.1. For a given screening rate s, define $s_{days} := T/s$, representing the number of days between two consecutive screening efforts, with $s = 0$ (no screening) yielding $s_{days} \to \infty$. Recall that $p(s)$ denotes the expected proportion of the infectious period (t_I) during which an unvaccinated and infected individual is in isolation, hence cannot transmit the infection. Consider an unvaccinated and infected subject (i.e., the following is conditional on the presence of the infection). The subject remains infectious for t_I days following exposure. Let T_s denote the time from exposure to the next screening test. Since screening occurs periodically every s_{days}, $T_s \sim U(0, s_{days})$. Let T_{nd} represent the duration during which the subject is infectious but undetected. Due to the assumptions of a perfectly reliable test and immediate isolation upon detection, we have

$$T_{nd} = \min\{T_s, t_I\}, \tag{S.19a}$$

$$p(s_{days}) = 1 - \frac{E[T_{nd}]}{t_I}. \tag{S.19b}$$

Case 1: $0 \le s_{days} \le t_I$: There is at least one screening during the infectious period t_I, hence the subject is guaranteed to be screened before the end of their infectious period, that is, $T_{nd} = \min\{T_s, t_I\} = T_s$. Then it readily follows that, $E[T_{nd}] = E[T_s] = \frac{s_{days}}{2}$. Substituting this expression into Equation (S.19b), we obtain

$$p(s_{days}) = 1 - \frac{s_{days}}{2t_I}.$$

Case 2: $t_I \le s_{days} \le T$: There is at most one screening during the infectious period t_I. Therefore,

$$\begin{aligned} E[T_{nd}] = E[\min\{T_s, t_I\}] &= \int_0^{t_I} t_s \cdot f_{T_s}(t_s) \cdot dt_s + \int_{t_I}^{s_{days}} t_I \cdot f_{T_s}(t_s) \cdot dt_s \\ &= \int_0^{t_I} t_s \cdot \frac{1}{s_{days}} \cdot dt_s + \int_{t_I}^{s_{days}} t_I \cdot \frac{1}{s_{days}} \cdot dt_s \\ &= t_I - \frac{t_I^2}{2s_{days}}. \end{aligned}$$

Substituting the above expression into Eq. (S.19b), we obtain

$$p(s_{days}) = \frac{t_I}{2s_{days}}.$$

Case 3: $s_{days} > T$: There is no screening during the semester ($s = 0$), hence $p(s_{days}) = 0$, and

$$p(s_{days}) = \begin{cases} 1 - \frac{s_{days}}{2t_I}, & \text{if } 0 \le s_{days} \le t_I \\ \frac{t_I}{2s_{days}}, & \text{if } t_I \le s_{days} \le T \\ 0, & \text{if } s_{days} > T \end{cases}. \tag{S.20}$$

Then, using the identity $s_{days} = T/s$, $p(\cdot)$ can be expressed as a function of the screening rate $s \in [0, T]$.

□

LEMMA 1. *For a given mitigation plan (d, s), cumulative infections at semester end, $I_T(d, s)$, follows:*

$$I_T(d,s) = \frac{K(d,s)}{1 + \left(\frac{K(d,s)-I_0}{I_0}\right) \cdot e^{-r(d,s)\cdot T}}, \tag{16}$$

where $K(d,s) = P \cdot (1 - cvg(d,s) \cdot \beta)$, $r(d,s) = \ln\left(R_0 \cdot [1 - p(s) - cvg(d,s) \cdot (\beta - p(s))]\right)/\tau$, and both parameters decrease in each of d and s.

Proof of Lemma 1. We begin by conditioning on vaccination status and applying the law of total probability to express R_e under a given intervention policy (d, s). In particular, the population is divided into two mutually exclusive groups: (1) vaccinated individuals (with proportion $cvg(d, s)$), each with a reduced probability of transmitting the infection, given by $1 - \beta$; (2) unvaccinated individuals (with proportion $1 - cvg(d, s)$), each isolated for a fraction $p(s)$ during their infectious period, hence each with a probability of transmitting the infection of $1 - p(s)$. Therefore, we have

$$\begin{aligned} R_e(d,s) &= R_0 \cdot [cvg(d,s) \cdot (1-\beta) + (1 - cvg(d,s)) \cdot (1 - p(s))] \\ &= R_0 \cdot [1 - p(s) - cvg(d,s) \cdot (\beta - p(s))]. \end{aligned} \tag{S.21}$$

Then, by Eq. (S.18),

$$r(d,s) = \frac{\ln(R_e(d,s))}{\tau} = \frac{ln\Big(R_0 \cdot [1 - p(s) - cvg(d,s) \cdot (\beta - p(s))]\Big)}{\tau}.$$

The carrying capacity $K(d, s)$ represents the number of susceptible individuals. Since each vaccinated individual with effective immunity is removed from the susceptible pool (which occurs with probability β), the adjusted carrying capacity is:

$$K(d,s) = P \cdot (1 - cvg(d,s) \cdot \beta). \tag{S.22}$$

Then the expression for I_T directly follows from Eq. (S.17). Finally, the result that parameters $r(d, s)$ and $K(d, s)$ decrease in d and s directly follow from Eqs. (S.18), (S.21), and (S.22), completing the proof. □

REMARK D.1. The expression $I_T(d, s)$ in Lemma 1 can be extended to the three-stage **QM-3** to account for both the initial screening rate and the subsequent adjustment at time t'. For a given mitigation plan (d, s, s'), $I_T(d, s, s')$ can be expressed as follows:

$$I_T(d,s,s') = \begin{cases} \dfrac{K(d,s)}{1 + \left(\dfrac{K(d,s) - I_{t'}\left(d, \frac{s \cdot t'}{T}\right)}{I_{t'}\left(d, \frac{s \cdot t'}{T}\right)}\right) \cdot e^{-r_2(d,s,s') \cdot (T - t')}} & if\, s \geq s', \\ \dfrac{K(d,s')}{1 + \left(\dfrac{K(d,s') - I_{t'}\left(d, \frac{s \cdot t'}{T}\right)}{I_{t'}\left(d, \frac{s \cdot t'}{T}\right)}\right) \cdot e^{-r_2(d,s,s') \cdot (T - t')}} & s < s', \end{cases} \tag{S.23}$$

where $r_2(d, s, s') = \ln\left(R_0 \cdot [1 - p(s') - cvg(d, \max\{s, s'\}) \cdot (\beta - p(s'))]\right)/\tau$.

Table C2: Parameter values and data sources used in the LGM numerical study.

Parameter (Definition)	Value	Range across Diseases	Data Source / Reference
I_0 (Initial number of infections)	5		
R_0 (Basic reproduction number)	5	[1.5,18]	Guerra et al. (2017), Muzembo et al. (2024)
β (Vaccine efficacy against infections)	0.8	[0.5,0.9]	Trombetta et al. (2022), Zeng et al. (2022)
τ (Generation time, days)	5	[3,11]	Xu et al. (2023), Sender et al. (2022)
t_I (Infectious period, days)	10	[3,16]	Hakki et al. (2022), Van Kampen et al. (2021)
T (Semester length, days)	80		

D.2. Details of the LGM Numerical Study (§5.2)

D.2.1. Base-case Parameters: Per-person budget $B = \$60$, per-test cost $\gamma = \$10$, minimum budget-compliance probability $\alpha = 0.9$. Epidemiology parameters for base-case and ranges for sensitivity analyses for the LGM numerical study (§5.2) are provided in Table C2.

D.2.2. Nature's Distributions and Parameters: Nature's vaccination disutility distribution, $f_W^N(\cdot)$, initial vaccination coverage, x_v^N, and screening disutility function, $h^N(\cdot)$, which are unknown to the planner (see §2.2), are constructed as follows.

In the base-case, $h^N(s) = s$ (with sensitivity analysis on nonlinear forms given by $h^N(s) = 15^{\rho} s^{1-\rho}$ for $\rho \in [-1.0, 0.9]$, see §5.2 and Figure 4). We model populations with varying vaccination characteristics through a Monte Carlo simulation with 5,000 replications. Each replication involves random draws for: x_v^N, sampled from $X_v^N \sim \mathcal{N}\big(\mu_{X_v^N} = 0.5,\, \sigma_{X_v^N} = 0.0239\big)$, and $f_W^N(\cdot)$, sampled from the ambiguity set $\mathcal{D} := \{Normal, LogNormal, Uniform, Gamma(shape), shape \in \{0.5, 1.0, 2.0, 3.0, 7.5\}\}$ (each distribution equally likely). (The degree of right-skewness decreases as the Gamma $shape$ parameter increases, with $shape = 1$ representing the exponential distribution). For each sampled x_v^N realization and $W^N(\cdot)$ distribution, we use the following procedure to calibrate the parameters of $f_W^N(.)$: first, we determine the corresponding parameters for the Normal distribution, designated as the *reference distribution*, such that the initial coverage satisfies $x_v^N = F_W^N(0)$ (Eq. (2)) and the standard deviation matches the planner's estimated σ_W. Then, depending on the number of free parameters requiring calibration (i.e., LogNormal requires three parameters, whereas all other distributions in the ambiguity set $\mathcal{D}$ require two parameters), we match the following information in the given order: (1) the value of x_v^N; (2) the shift amount, obtained from the 5$^{\text{th}}$ percentile of the reference distribution; and (3) σ_W.

D.2.3. Planner's Empirical Estimation and Calibration: The planner conducts a pre-semester survey by randomly sampling 10% of the campus population (2,400 participants), each representing a random draw w from Nature's distribution $f_W^N(\cdot)$. Based on survey responses, the planner estimates an empirical distribution $f_W(\cdot)$ via Maximum Likelihood Estimation (MLE) from candidate distributions in ambiguity set $\mathcal{D}$. Consistent with the analytical framework (§3.1), the planner maintains uncertainty on initial vaccination coverage X_v, where each realization x_v provides a perfect signal for the conditional distribution

$f_{W;x_v}(\cdot)$ (see §3.1). The planner uses historical or external vaccination data to construct the prior, $X_v \sim$ Normal$(\mu_{X_v}, \sigma_{X_v^N})$, with μ_{X_v} potentially inaccurate, but $\sigma_{X_v^N}$ assumed accurate (robustness checks with inaccurate σ_{X_v} yielded similar insights). To construct the set $\{f_{W;x_v}(\cdot), \forall x_v \in [0,1]\}$, the planner models that each $f_{W;x_v}(\cdot)$ is a *shifted* form of the empirical distribution $f_W(\cdot)$, satisfying $F_{W;x_v}(0) = x_v,\ \forall x_v \in [0,1]$. This calibration procedure ensures consistency between prior beliefs, survey data, and assumptions of the optimization models, while maintaining numerical tractability. In the *no-survey* benchmark, the planner assumes $W \sim Normal$. The planner models the random screening disutility function as $H(s) = \mathcal{K}s$, where $\mathcal{K} \sim U(0, 2 \cdot \mu_{\mathcal{K}})$, with prior mean $\mu_{\mathcal{K}}$.

D.2.4. Constructing the Pareto Sets and Identifying the Optimal Strategies: We use a numerical optimization framework in Python to generate the Pareto set, and identify the optimal strategy for a given objective function for each model and scenario. The structural results from Propositions 1 and 2, Corollaries 1 and 2 and Remark A.2 play an important practical role in this implementation by characterizing the Pareto set for each model, thereby restricting the search space to the Pareto set. Generating the Pareto set involves solving the root-finding problem repeatedly to determine the respective s- and d-roots, which is implemented via the Nelder-Mead Algorithm (Gao and Han 2012).

Further, to identify the optimal strategy that maximizes the number of infections averted for each model and scenario, we evaluate the expected number of infections averted over all policies in the Pareto set: (d, s) for **JM** and **QM-2**, and (d, s, s') for **QM-3**. For each candidate, we compute $\mathcal{G}(d, s; x_v)$ (and $\mathcal{G}(d, s, s'; x_v)$ for **QM-3**) using the calibrated LGM (see §5.1, §D.2, and §E.2), and then take expectations wrt X_v using the distribution specified in Appendix D.2.3.

Appendix E: Details of the Case Study (§6) and Figures

E.1. Data Sources and Model Calibration

In the case study, we assume that the planner has perfect knowledge of Nature's initial vaccination coverage (x_v^N) and screening disutility function ($h^N(\cdot)$); their calibration is described next.

E.1.1. Fitting a Distribution for Nature's X_v^N (Initial Vaccination Coverage): Based on the the county level data for the states of AL and CO (CDC 2023a) on December 18, 2021 (a representative date in the planning period), we fit a distribution for X_v^N (initial vaccination (boosted) coverage) for each state by utilizing the @RISK software (Lumivero 2023): using several goodness-of-fit tests and measures (Kenton 2023) including chi-square tests, the Akaike information criterion, and Q-Q plots, @RISK fits the distribution of X_v^N as $Normal(0.25, 0.028)$ and $Normal(0.42, 0.065)$ for AL and CO, resp.. Due to the unavailability of UA's initial coverage data, we utilized an alternative estimation method. Based on UA's 2023 undergraduate enrollment data (College Factual 2023), in-state and out-of-state proportions are 32.8% and 67.2% (predominantly from eleven states), resp.. Then, for AL and each of these eleven states, we use the proportion of the state's 18-24 year old population vaccinated with at least one dose by August 5, 2021 (CDC 2023b)

Table C3: Parameter values used in the case study.

$X_v^{N\dagger} \sim Truncated\,Normal$	UA: $\mu_{X_v}^N = 0.295, \sigma_{X_v}^N = 0.021$ AL: $\mu_{X_v}^N = 0.25, \sigma_{X_v}^N = 0.028$ CO: $\mu_{X_v}^N = 0.42, \sigma_{X_v}^N = 0.065$
Objectives‡ (health reward)	Infections averted Hospitalizations averted
$h^N(s)$	s, $2s$, $3s$, $5s$
T (Semester length, days)	80
P (Campus population)	24,000 (22,500 students, 1,500 faculty)
I_0 (Initial number of infections)	48 (0.2% of total population)
R_0 (Basic reproduction number)	Compartmental model: 12 (students), 6.4 (faculty) LGM: 11.65 (weighted average over the population)
β (Vaccine efficacy against infections)	0.7805
Vaccine efficacy against hospitalizations	0.9525
Vaccine efficacy against fatality	0.97155
ρ (Recovery rate)	0.0667 per 8-hour cycle
t_I (Infectious period, i.e., $1/\rho$, days)	5
Latent period (days)	3.5
τ (Generation time, i.e., latent period + infectious period, days)	8.5
Rate of symptom onset	0.0286 per 8-hour cycle
Rate at which exposed subjects become asymptomatic and infectious	0.095 per 8-hour cycle
Test accuracy	Sensitivity: 80%, Specificity: 98%

† $x_v^N \in [0,1]$ is discretized in increments of 0.004.
‡ $d \in [0, d_{root}(0; x_v^N)]$ and $s \in [0, 80]$ are discretized to create 100 points each.

(a representative date for the start of UA's Fall 2021 semester), as a proxy for the state's initial coverage. A weighted average (based on UA's enrollment data) of the initial coverage values for AL and the top eleven states yields UA's initial coverage of 29.5% with standard deviation of 2.1%.

To ensure that $X_v^N \in [0,1]$, we truncate each Normal distribution between 0 and 1. Because the standard deviation values of the fitted distributions are relatively small, truncation does not affect the distribution parameters within three decimal point accuracy. Hence, we model X_v^N as $Truncated\,Normal(0.25, 0.028)$ for AL, $Truncated\,Normal(0.42, 0.065)$ for CO, and $Truncated\,Normal(0.295, 0.021)$ for UA.

E.1.2. Fitting a Distribution for Nature's W (Vaccination Disutility): Ideally, the distribution of nature's W would be constructed from data on individuals' disutility towards (alternatively, the monetary reward needed for) vaccination, obtainable, e.g., via surveys. In the absence of such data in the existing literature, we rely on data from UA's experience for illustrative purposes, supplemented by sensitivity analysis. UA offered a one-time incentive of $\$40$ for each student vaccinated at the beginning of the Fall 2021 semester, resulting in a (final) vaccination coverage of around 61% among UA's full-time student body (Jones 2023). Combined with our estimated initial coverage, $X_v^N \sim Normal(0.295, 0.21)$, the observed coverage of 61% at $\$40$ provides two reference points for calibrating the distribution of nature's W.

Nature's W is generated through the following procedure. First, a distribution family is randomly selected from the ambiguity set $\mathcal{D}$. Next, a realization of the initial coverage x_v^N is drawn. Finally, the parameters of the sampled distribution are calibrated so that $F_W^N(0) = x_v^N$ and $F_W^N(40) = 0.61$. This calibration is

feasible for all families in $\mathcal{D}$ except the LogNormal, since these families have two free parameters that can be uniquely determined by the imposed conditions. In the case of the LogNormal distribution, which involves three free parameters (including a location or shift parameter), a modified procedure is applied: The unconditional distribution of W is constructed such that $F_W^N(0) = \mu_{X_v^N}$ and $F_W^N(40) = 0.61$. Then, for any realization x_v^N, the associated conditional distribution is obtained by shifting the unconditional distribution so that $F_W^N(0) = x_v^N$.

We extend the analysis to the AL and CO settings, where only the initial vaccination coverage is available, and coverage at an alternative incentive level (e.g., \$40) is not observed. To address this limitation, we assume AL and CO exhibit the same proportional responsiveness to incentives as observed for UA; then the coverage at \$40 is estimated as $F_W^N(40) = \mu_{X_v}^N \cdot \frac{0.61}{0.295}$, yielding values of 0.517 for AL and 0.868 for CO.

E.1.3. Fitting the Nature's $h^N(\cdot)$ Function (Disutility Towards Routine Screening): Ideally, function $h^N(s)$, which quantifies the screening disutility in monetary terms, would be derived from empirical data reflecting individuals' disutility from undergoing screening at various frequencies, potentially gathered via surveys. Given the lack of such specific data in the existing literature, we consider a linear function, $h^N(s) = \min\{ks, M\}$, bounded by a sufficiently large $M \geq \overline{S}$ in accordance with Assumption 1 (§2.1); for notational simplicity, the bound M is omitted in $h^N(s)$ expressions. For the base-case, we consider $k = 1$, and perform sensitivity analysis over $k \in \{2, 3, 5\}$ (Table C3), see Figures 6 and 7 and Appendix E.4 for the results.

E.1.4. Overview of the Compartmental Model used for Evaluation: To evaluate the mitigation policies produced by the LGM-embedded optimization framework, we use a medically detailed deterministic compartmental model developed in prior work (Rabil et al. 2022b) and extended in this paper to incorporate semester- long vaccination uptake under incentive-based policies. We summarize the key features below and refer the interested reader to Rabil et al. (2022b) for the full model.

The model is an extended SEIR-type framework tailored to COVID-19 mitigation on a college campus during Spring 2022; the planning occurs in December 2021. It captures heterogeneity across two population *groups* (students and faculty/staff), three *vaccination states* (unvaccinated, vaccinated and not boosted with waning immunity (hereafter vaccinated), and vaccinated and boosted (hereafter boosted)), and two *co-circulating variants* (Delta and Omicron) at equal proportions. Disease transmission, hospitalization risk, and fatality risk depend on both group and vaccination status; the model also incorporates externally imported infections (I_0) to campus.

The population is partitioned into four main pools of compartments. The *transmission pool* contains susceptible, exposed, and asymptomatic infectious individuals, indexed by group and vaccination status. The *recovered pool* contains individuals with infection-induced immunity, distinguished by whether prior infection was known versus unknown by the university. The *isolation/hospital pool* contains false-positive, true-positive asymptomatic, infected symptomatic, and hospitalized individuals; subjects in this pool do not

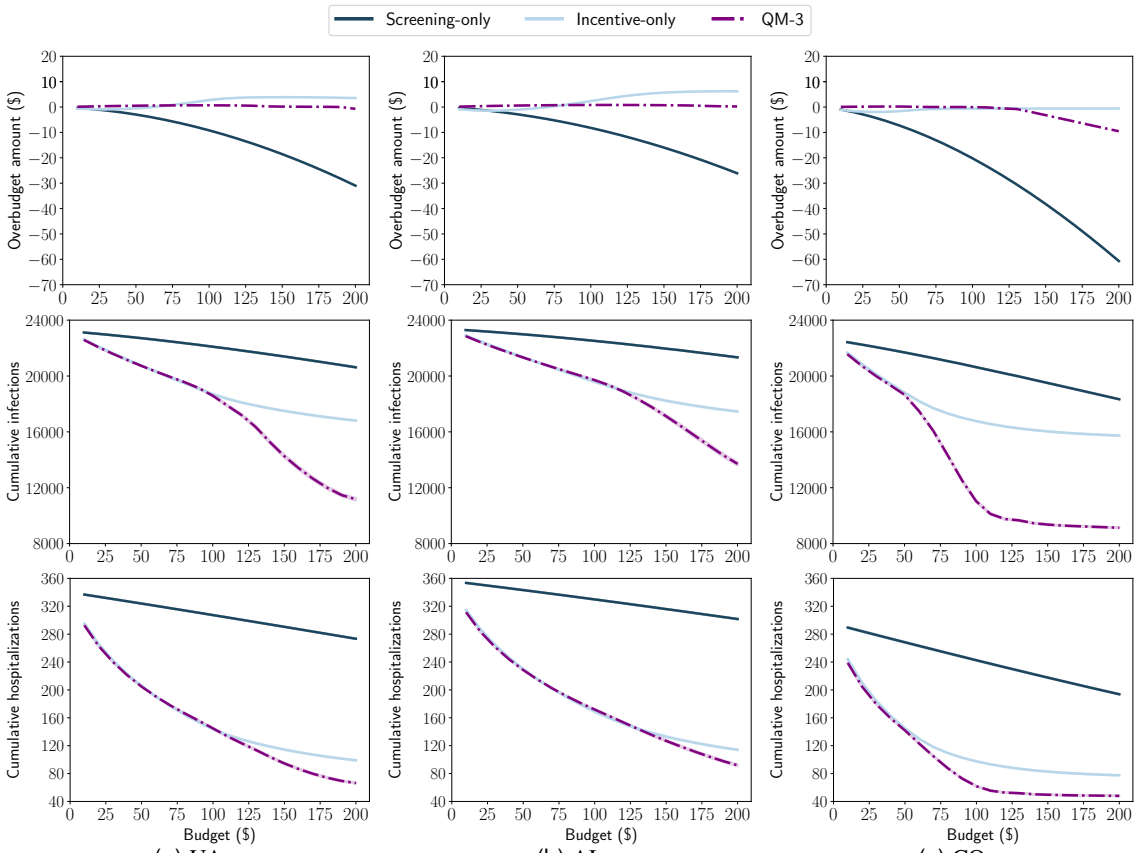


Figure C1 Overbudget ($), cumulative infections, and cumulative hospitalizations for **QM-3**, and **Screening-only** and **Incentive-only** benchmarks for UA, AL, and CO scenarios, and Nature's screening disutility function, $h^N(s) = s$. Vaccinations are assumed to be uniformly spread over the semester.

contribute to transmission. Finally, the *removed pool* contains deaths. The model tracks 62 compartments in total, representing differential transmission and clinical outcomes across groups while preserving a tractable deterministic structure.

The model evolves in 8-hour cycles over the 80-day semester. Routine screening is applied to the eligible screening pool. Test sensitivity and specificity are not perfect; test results are available in 8 hours; false-positives are temporarily isolated and released the next day; and true-positive or symptomatic individuals remain isolated until recovery or hospitalization. Isolation is assumed to prevent all future transmissions.

The model distinguishes vaccine- and infection-induced immunity. Individuals may begin the semester unvaccinated, vaccinated, or boosted, and may also acquire natural immunity after recovery from infection. Vaccine efficacy against infection and hospitalization is variant-dependent, and waning immunity is incorporated only for individuals vaccinated before the semester but not recently boosted. In the extended version developed in this paper, additional vaccinations induced by incentives or screening may occur before or during the semester, depending on the scenario considered.

For each mitigation policy, the compartmental model provides the cumulative infections, hospitalizations, and total screening tests over the semester. We use these outputs as the evaluation layer of the case study,

allowing us to assess how the LGM-embedded optimization models perform under a richer epidemic model with heterogeneous transmission, hospitalization risk, and fatality risk, imperfect test, delayed isolation, clinical progression, external shocks, and continous vaccination.

E.1.5. Additional Case Study Parameters and Implementation Details: All financial parameters are given in US dollars. While the testing cost (γ) depends on the setting, it is primarily the ratio, B/γ, that drives an optimal strategy. Therefore, setting $\gamma = 10$, we vary the per-person budget in $B \in (0, 200]$, where the budget upper bound allows the screening of the entire population at the maximum rate were there no monetary incentives. The UA, AL, and CO base-cases consider the objective of averting infections and a minimum budget-compliance probability of $\alpha = 0.90$. Optimal strategies in the case study are determined using the same numerical approach described in Appendix D.2.4 (using the distribution specified in Table C3), where the structural results identify the Pareto set for each model, and the candidate policies in the Pareto set are evaluated under the case study objective, i.e., maximizing infections averted.

E.2. Planner's LGM Calibration for the Case Study and Robustness to Misspecification

The LGM-embedded optimization model used by the planner in the case study is calibrated from the same epidemiological setting as the compartmental model used for evaluation (see Table C3), but serves a different purpose. The compartmental model governs Nature's underlying disease dynamics, whereas the LGM serves as the planner's tractable forecasting model for mitigation policy optimization. Accordingly, the key epidemiological inputs for the LGM are chosen to remain aligned with the case study setting, aggregating where necessary to maintain a single-population representation.

In particular, the planner's base-case values for I_0, τ, t_I, and β are consistent among the two models. However, regarding the basic reproduction number (R_0), the LGM uses an aggregated value that assumes a homogeneous population of size P. Specifically, the compartmental model treats students and faculty as distinct groups, with populations of $22{,}500$ and $1{,}500$, and corresponding reproduction numbers of $R_0 = 12$ and $R_0 = 6.4$, respectively. By contrast, the planner's LGM consolidates these groups into a single homogeneous population of size $24{,}000$, with a weighted average reproduction number of $R_0 = 11.65$.

Because these LGM inputs must be estimated by the planner at the time of policy design, we evaluate model robustness by considering *misspecification* in the LGM epidemiological parameters. To this end, we hold Nature's compartmental model parameters fixed at the base-case values in Table C3, and vary, one at a time, the five epidemiological inputs that directly inform the planner's LGM (see §5.1 and §E.2): R_0, β, τ, t_I, and I_0. For each misspecified value, the planner re-optimizes the LGM-embedded **QM-3**; the resulting policy is then evaluated under Nature's true compartmental dynamics and compared with the **Incentive-only** and **Screening-only** benchmarks to determine whether the gains of **QM-3** persist under forecast error.

The misspecification ranges are centered on late-2021/early-2022 evidence relevant to the Spring 2022 Omicron/Delta transition. For parameters with published interval estimates, we use literature-informed ranges, summarized in Table C4, that are motivated by the corresponding 95% confidence intervals.

Table C4: Planner-side robustness ranges used in the model-misspecification analysis in the case study.

Parameter (Definition)	Robustness Range	Data Source / Reference
I_0 (Initial number of infections)	[10,100]	
R_0 (Basic reproduction number)	[8,14]	Ito et al. (2022), Khandia et al. (2022)
β (Vaccine efficacy against infections)	[0.70,0.85]	Andrews et al. (2022), Abu-Raddad et al. (2022)
τ (Generation time, days)	[7,10]	Manica et al. (2022), Park et al. (2023)
t_I (Infectious period, days)	[4,6]	Boucau et al. (2022), Wu et al. (2023)

Figures C4 and C5 summarize cumulative infections and hospitalizations under **QM-3** relative to the **Incentive-only** and **Screening-only** benchmarks for $B = \$150$, a moderate budget level consistent with the range considered in the case study. Figure C4 shows that under the base screening-disutility ($h(s) = s$), **QM-3** consistently outperforms both benchmarks in cumulative infections across the full misspecification range for all five parameters. The largest variation arises under the misspecification of R_0, while misspecification of t_I and I_0 has little effect on the realized infection count. Across the full misspecification range, cumulative infections under **QM-3** vary from approximately $8,100$ to $10,100$. Hospitalizations exhibit a similar qualitative robustness in nearly all cases, but are more sensitive to misspecification in R_0 because hospitalization outcomes in this setting depend strongly on vaccination-induced protection. As shown in Figure C5, **QM-3** continues to dominate both benchmarks when the planner misspecifies β, τ, t_I, or I_0. Under significantly underestimated values of R_0, however, **QM-3** can slightly underperform the **Incentive-only** benchmark. This pattern is intuitive: when the planner perceives transmission to be mild, the LGM-embedded optimization model shifts the policy toward heavier screening (see Figure 5a) and lower incentive spending, whereas hospitalizations remain highly sensitive to vaccine coverage. As the planner's R_0 moves back toward the base-case region, **QM-3** regains a clear advantage in the number of hospitalizations.

Similar conclusion holds in other scenarios, omitted due to space limitations. For $B = 100, k = 1$, cumulative infections and hospitalizations under **QM-3** remain within $[12{,}991, 13{,}497]$ and $[70.5, 71.9]$, resp., across all misspecification scenarios. For $B = 100, k = 5$, the corresponding ranges are $[5{,}104, 5{,}501]$ and $[7.05, 8.64]$, and for $B = 150, k = 5$, they are $[5{,}036, 5{,}292]$ and $[6.44, 8.43]$. Thus, while the LGM sensitivity analysis in §5.2 shows that the planner's optimal strategy can shift as these epidemiological inputs vary (Figure 5), the outcomes evaluated by the compartmental model are substantially more stable. The reason is that, in the Spring 2022 campus setting, many such shifts occur among policies that induce similar vaccination coverage, the dominant driver of both infections and hospitalizations, so the health gains of **QM-3** over the benchmarks remain robust to moderate planner misspecification.

E.3. Vaccination Throughout the Semester

In this section we examine the base-case under a scenario where vaccinations occur uniformly throughout the semester. We adapt **QM-3** to account for the uniform vaccination dynamics during the semester and adjust decisions (d, s, s') accordingly. Figure C1 provides the overbudget amount, cumulative infections,

and cumulative hospitalizations for **QM-3**, and **Incentive-only** and **Screening-only** benchmarks: **QM-3** consistently outperforms both benchmarks in terms of budget adherence and cumulative infection reduction across the UA, AL, and CO scenarios, even when vaccinations occur uniformly throughout the semester.

E.4. Additional Figures

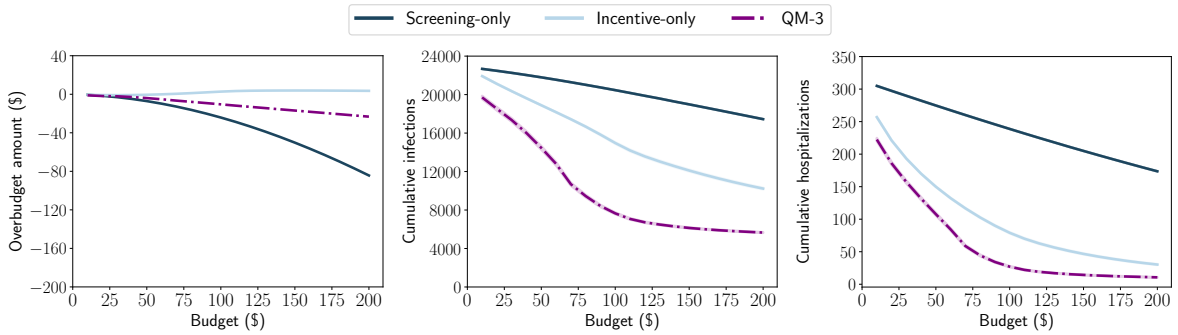


Figure C2 Overbudget ($), cumulative infections, and cumulative hospitalizations for **QM-3**, and **Screening-only** and **Incentive-only** benchmarks for UA scenario, and Nature's screening disutility function, $h^N(s) = 2s$.

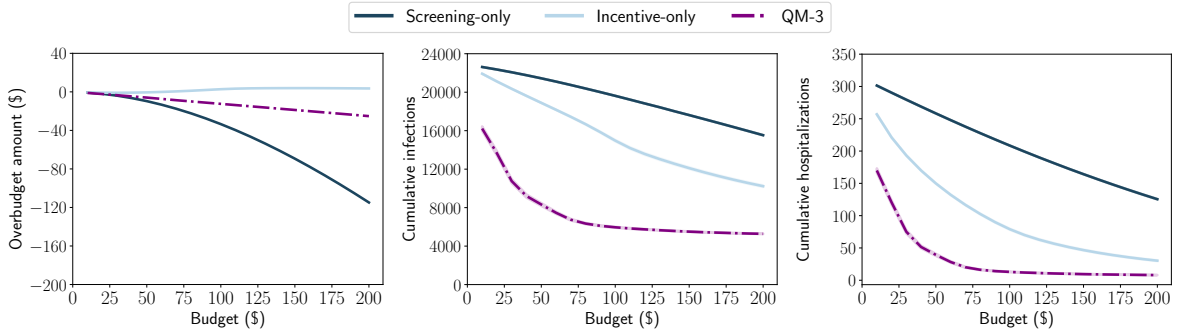
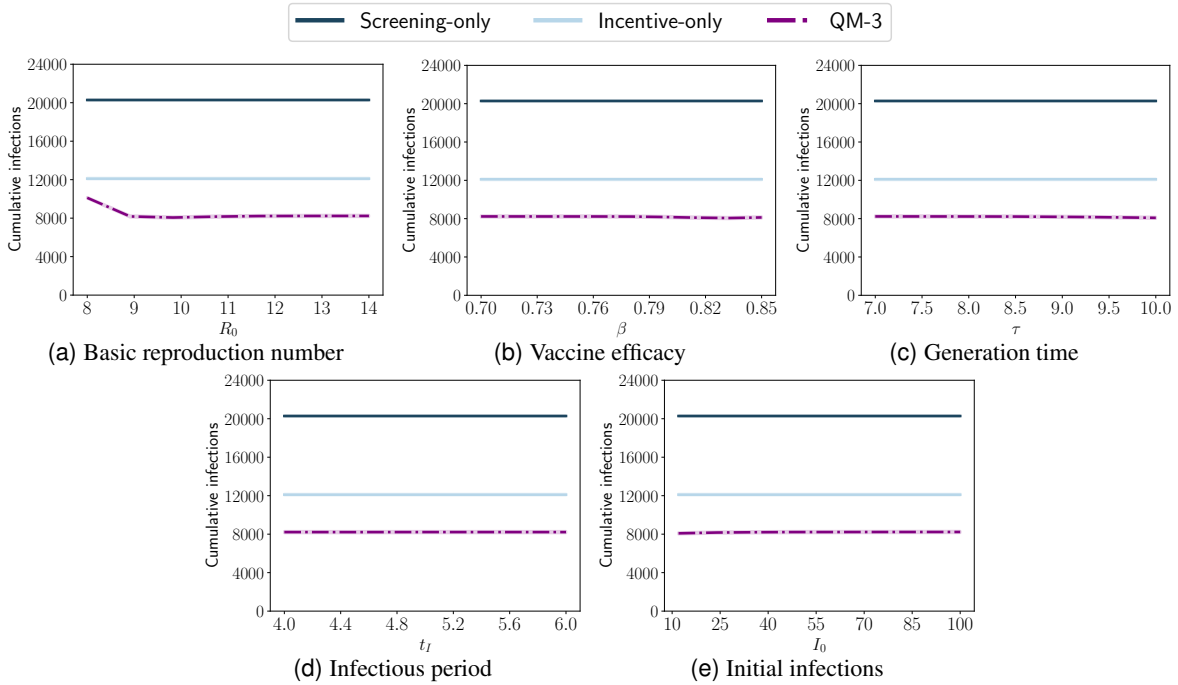


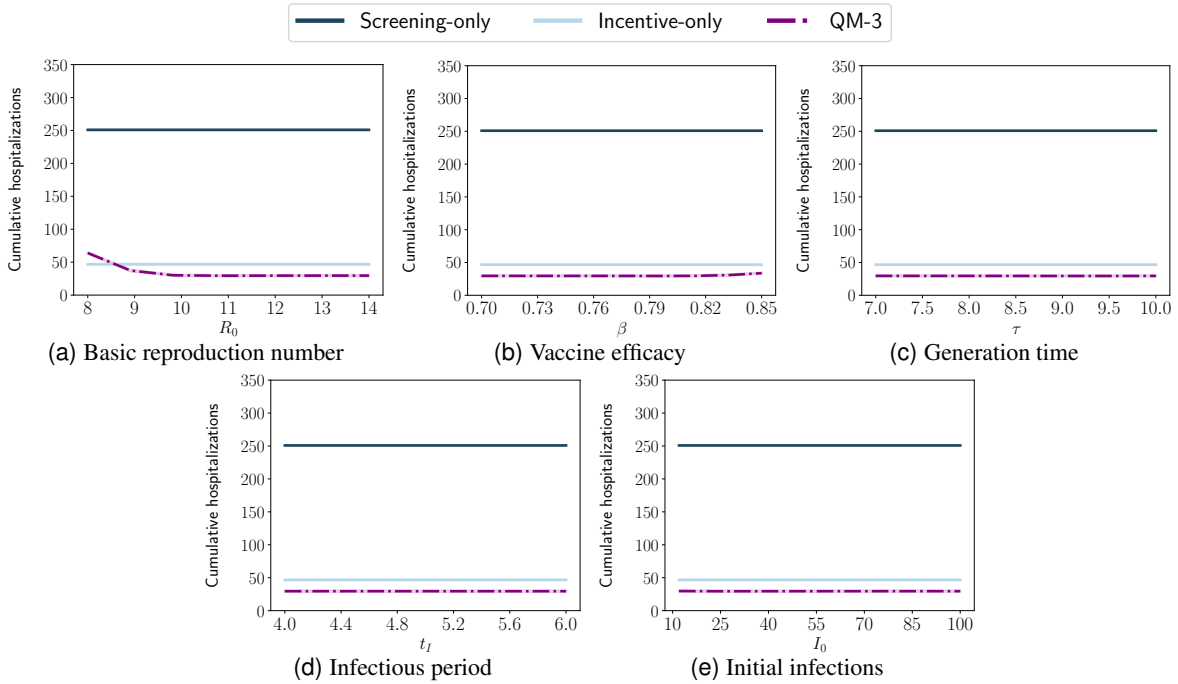
Figure C3 Overbudget ($), cumulative infections, and cumulative hospitalizations for **QM-3**, and **Screening-only** and **Incentive-only** benchmarks for UA scenario, and Nature's screening disutility function, $h^N(s) = 3s$.



(a) Basic reproduction number (b) Vaccine efficacy (c) Generation time

(d) Infectious period (e) Initial infections

Figure C4 Cumulative infections for **QM-3**, **Incentive-only**, and **Screening-only** under the planner's misspecification of a single epidemiological input in LGM-embedded optimization ($R_0 \in [8, 14]$, $\beta \in [0.70, 0.85]$, $\tau \in [7, 10]$, $t_I \in [4, 6]$, $I_0 \in [10, 100]$); Nature uses base-case parameterization (Table C3). UA scenario, $B = \$150$, $k = 1$.



(a) Basic reproduction number (b) Vaccine efficacy (c) Generation time

(d) Infectious period (e) Initial infections

Figure C5 Cumulative hospitalizations for **QM-3**, **Incentive-only**, and **Screening-only** under the planner's misspecification of a single epidemiological input in LGM-embedded optimization ($R_0 \in [8, 14]$, $\beta \in [0.70, 0.85]$, $\tau \in [7, 10]$, $t_I \in [4, 6]$, $I_0 \in [10, 100]$); Nature uses base-case parameterization (Table C3). UA scenario, $B = \$150$, $k = 1$.