medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

# Optimal epidemic control under uncertainty: trade-offs between information collection and other actions

Julien Flaig* Nicolas Houy†

March 28, 2022

**Abstract**

Background: Recent epidemics and measures taken to control them – through vaccination or other actions – have highlighted the role and importance of uncertainty in public health. There is generally a trade-off between information collection and other uses of resources. Explicitly or more implicitly, the concept of expected value of perfect information (EVPI) is central in order to inform policy makers in an uncertain environment.

Method: We use a simple SIR disease emergence and transmission model with vaccination that can be administered as one or two doses. The disease parameters and vaccine characteristics are uncertain. We study the trade-offs between information acquisition and two other measures: bringing vaccination forward, and acquiring more vaccine doses. To do this, we quantify the EVPI under different constraints faced by public health authorities, *i.e.* the time of the vaccination campaign implementation and the number of vaccine doses available.

Results: We discuss the appropriateness of different responses under uncertainty. We show that in some cases, vaccinating later or with less vaccine doses but more information may bring better results than vaccinating earlier or with more doses and less information respectively.

Conclusion: In the present methodological paper, we show in an abstract setting how clearly defining and treating the trade-off between information acquisition and the relaxation of constraints can improve public health decision making.

**Keywords:** value of information, EVPI, emerging infectious disease, epidemic control, uncertain environment.

*Epidemiology and Modelling in Infectious Diseases (EPIMOD), F-01240 Lent, France. Email: jflaig@epimod.fr

†University of Lyon, Lyon, F-69007, France; CNRS, GATE Lyon Saint-Etienne, F-69130, France. Email: houy@gate.cnrs.fr.



NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

# 1 Introduction

Infectious disease control is steeped in uncertainty. Uncertainty may concern the natural history of the pathogen, the successful development of prophylactic or curative treatments, the ability to deliver control interventions, or the population characteristics, including contact patterns and individual behaviors [1]. Since the spread of an infectious disease is a dynamic and nonlinear phenomenon, uncertainties can seriously hinder control efforts. Control measures taken at a certain time given available information may prove sub-optimal as more information is gathered about the epidemic, but these measures are usually irreversible and their cost is sunk.

Yet learning about an epidemic costs both money and time, so there is a trade-off between gaining information and allocating these resources to other uses. Money, of course, can be used to produce medical supplies or build up intervention capacity in general. As for time, it may seem sensible to act earlier rather than later to minimize the impact of an epidemic, insofar as it can spread as time goes by. How, then, to balance the benefit of gaining information against the benefit of alternative resource uses? This can be done through value of information analyses, which allow to quantify the benefit of reducing uncertainties.

Value of information analyses are widely used in the fields of public health and health economics [2, 3]. One popular value of information metric is the expected value of perfect information (EVPI), which is computed as follow. Put generally, a decision-maker chooses a strategy $a$ in a set $A$ of alternatives to maximize a value function $V$ (e.g. the number of averted cases or averted costs over the course of the epidemic) that also depends on uncertain parameters denoted $\xi$. Under uncertainty about the true value of $\xi$, an alternative $a_0$ is usually picked that maximizes the expected value over possible realizations of $\xi$: $a_0 = \arg\max_{a \in A} E_\xi \left[V(a, \xi)\right]$. If there is no uncertainty and the true value $\xi^*$ of $\xi$ is known, an alternative $a^*$ can be picked that simply maximizes the value $V$ given $\xi^*$: $a^* = \arg\max_{a \in A} V(a, \xi^*)$. The difference $V(a^*, \xi^*) - V(a_0, \xi^*) = \max_{a \in A} V(a, \xi^*) - V(a_0, \xi^*)$ is the value loss due to making a decision under uncertainty. EVPI is the expected value loss over possible realizations of $\xi$: $E_\xi \left[\max_{a \in A} V(a, \xi) - V(a_0, \xi)\right] = E_\xi \left[\max_{a \in A} V(a, \xi)\right] - \max_{a \in A} E_\xi \left[V(a, \xi)\right]$. In other words, EVPI quantifies uncertainty as the expected benefit of resolving it, and hence gives an upper bound on the cost one should be willing to pay to reduce it.

A simple tweak allows to use an EVPI framework to balance information acquisition against other resource uses. The two terms of EVPI can be computed separately:

1. the *ex ante* maximum expected value $\max_{a \in A} E_\xi \left[V(a, \xi)\right]$, that is the expected value of

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

controlling the epidemic without collecting more information (i.e. *before* the true value of $\xi$ is known), and

2. the expected *ex post* maximum value $E_\xi \left[\max_{a \in A} V(a, \xi)\right]$, that is the expected value of controlling the epidemic with perfect knowledge of $\xi$ (i.e. *after* the true value of $\xi$ is known).

Using resources for other uses than collecting information may, for instance, broaden the set $A$ of feasible policies. Thus, computing the two terms of the EVPI for different sets $A$ of feasible policies allows to explore the trade-off between collecting information and alternative uses of resources. More generally, these alternative uses of resources could also improve the performance of a previously feasible policy. For clarity, in the present paper, we formulate "the same policy, but more efficient" as "another policy" and discuss all results in terms of the set of feasible policies.

Importantly, EVPI depends on modeling choices, some of which involve not only scientific but also political and ethical considerations. Determining the possible values of uncertain parameters and their likelihood is primarily a scientific or technical issue. The choice of a policy performance metric, by contrast, is also a political and ethical question: consider for instance the difference between maximizing the number of avoided deaths and saving a maximum of life-years. The same applies to the chosen objective. In the previous computations, we assumed that the objective is to maximise the expected value over uncertain parameter values. The value of information framework can be adapted to alternative objectives such as minimizing the probability of extreme adverse events (see the examples provided by Shea et al. [4]).

Policy-oriented studies in the field of infectious disease control rarely consider the value of acquiring more information (see [5] for an example of study which does), and considering information collection as a strategic choice among others is even less common. Most studies focus on available information only and disregard the possibility of resolving uncertainties. A typical approach consists in using available data to fit a disease transmission model. This task can be resource intensive and is the main contribution of many studies. It results in estimates of the uncertain parameters (point estimates or distribution estimates depending on the method used). In a second step, these estimates are used to determine the best policy either by comparing all relevant options exhaustively or via optimization methods. Finally, the robustness of the chosen policy is tested by performing sensitivity analyses. The objective of these sensitivity analyses is not so much to quantify uncertainties (e.g. by using value of information metrics to estimate the benefit of reducing them) as to justify the choice of a policy based on available data and the corresponding model fit. Available

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

information is assumed to be fixed and is used both to pick a policy, and to derive scenarios in which to test the performance of that policy.[1] Even considering the objective of justifying a policy given available information, sensitivity analyses can easily be incomplete (limited sampling of the parameter space) or use misleading metrics (e.g. the probability that a given policy will be the best instead of the expected value of that policy). We refer to Houy and Flaig [6] for a discussion of these issues and examples from the field of hospital epidemiology.

Other studies assume that more information is acquired over time and propose algorithms to estimate parameters (see [7] and the articles reviewed therein) or adapt interventions (see [8, 9], for instance) in real-time. Yet in these studies just as in studies assuming fixed information, information acquisition is seen as a *passive* process, and is not stemming from *active* strategic decisions. Namely, there is no trade-off between data collection and other uses of resources. Data is simply assumed to become available "for free" over time, and the point of these articles is to show how to use it. In this respect, whether information is assumed to be available at one point in time or to become available over time makes little difference.

To some extent, testing or screening optimization could be seen as strategic (or active) acquisition of information, however not in the sense of solving the trade-off between collection of information and alternative uses of resources as it does not entail assessing the value of information (at least not explicitly). In the context of infectious disease control, the objective of testing optimization is usually to control the spread of a disease by finding and isolating infectious individuals [10–14] or, less frequently, by finding recovered and immune individuals to end their isolation [15, 16]. Framing the problem in terms of value of information, the uncertain parameters $\xi$ would be the health statuses (susceptible, infectious, or recovered) of individuals. Yet the decision $a^*$ made once this information is available is typically predefined and fixed: it often simply consists in, say, isolating tested individuals found to be infectious, or in following a more sophisticated threshold-based rule by isolating a tested individual based also on other individuals' test results. In both cases, the problem is to find *how* to test (who and when) given that such or such predefined decision (isolate or not) rule is applied depending on test results, while a value of information analysis would rather address the issue of *whether* to test or not given that test information could be used to make better decisions.

In the present conceptual and methodological article, we use a generic deterministic

[1] This amounts to picking $a_0 = \arg\max_{a \in A} E_\xi [V(a, \xi)]$, and then testing the robustness of $a_0$ by computing $E_\xi [V(a_0, \xi)]$. Contrast this with value of information approaches: $a_0$ is deemed robust if the value $V$ cannot be much improved by knowing the true value $\xi^*$ of $\xi$.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

SIR infectious disease transmission model to make explicit and illustrate the trade-off between collecting information about an epidemic, and using resources to control it without collecting more information. The control intervention is vaccination, and the value $V$ of a vaccination policy is the number of averted symptomatic individual-days over the course of the epidemic compared to the no intervention scenario. We want to illustrate, first, how vaccinating earlier with less information compares to vaccinating later with more information and, second, how vaccinating with more vaccine doses but less information compares to vaccinating with less doses and more information. To do this, we compute the ex ante maximum value and and the expected ex post maximum value under different time (intervention date) and logistical (available vaccine doses) constraints. This allows to illustrate graphically the trade-off between relaxing these constraints and resolving uncertainties about the epidemic.

Our work is related to several previously published studies. Following a similar line of thought, some authors proposed multi-stage control strategies allowing to actively include information collection in decision making rather than passively adapt to new information. In the first example developed in their study, Shea et al. [4] showed how anticipating that uncertainty will be resolved at a later stage can influence decisions in earlier stages. The trade-off between acting early or collecting information and acting later is solved implicitly, but it is not the main focus of the article – the trade-off is solved for a single late (informed) intervention date. Similarly, Atkins et al. [17] considered a scenario in which the efficacy of a vaccine is uncertain but can be learned through vaccination campaigns. In this study as well, the focus is on the comparison of active information collection versus passive adaptation, and trade-offs are solved implicitly. The trade-off between trying to control an epidemic early with less information and later with more information is treated explicitly and in more details by Thompson et al. [18]. However they did not used an EVPI framework but proposed an heuristic algorithm allowing to anticipate future learning. We want to bring the discussion to a more general and elemental level. The relationship between logistical constraints (or equivalently resource or budget constraints) and the value of information was investigated by Shea et al. [4] (second example of the paper) and by Li et al. [19], however not in terms of the trade-off between relaxing these constraints and collecting information about the epidemic. This trade-off is illustrated by Woods et al. [20] in the case of HIV, but not for the full range of resource constraints. In this article, we address both implementation time and resource constraints and focus on trade-offs more in-depth.

The remainder of the paper is structured as follows. We introduce the generic disease transmission model and the control interventions to be used in simulations in Section 2.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

The simulation results are shows and discussed in Section 3. Section 4 concludes.

# 2 Materials and methods

We consider the emergence of an infectious disease in a closed population of $N = 10{,}000$ individuals. The spread of the disease is described by a SIR model with symptomatic and asymptomatic infections, and vaccination. Individuals may receive 0, 1, or 2 doses of vaccine. Vaccination is assumed to reduce infectiousness, susceptibility, and the probability of symptomatic infection. For concision and illustrative purposes, we use a deterministic model and leave aside stochastic effects, although they can be critical in the context of infectious disease emergence (see [18]). At time 0, no individual is vaccinated and 1/10,000 of the total number of individuals is assumed to be infected. Simulations are run until i) all vaccination interventions have taken place, and ii) the epidemic is extinct. In our continuous framework, we define that the epidemic is extinct when the number of infectious individuals is less than $N \times 10^{-8} = 10^{-4}$ individuals.

We assume the model structure to be known and focus on parameter uncertainties. The model equations are given in Appendix A and the model parameters in Table 1. The unit of time is the day throughout the article. We assume prior distributions to be available for all uncertain parameters. A set of prior distributions for parameters (PDP) represents i) the *level of information* about parameter values, and ii) *expectations* about these values at the time of decision. In the main text, we focus on PDP0, which assumes that some information is available about the parameters and that single dose and double dose vaccination are expected to provide intermediate protection. The details of PDP0 are given in Appendix B and Figure 1 shows 10,000 draws from PDP0.[2] We also provide results for two other PDPs in appendix. The first, PDP1, assumes a low level of information about the parameters. The second, PDP2, assumes that some information is available and that two doses of vaccine are expected to offer markedly more protection than a single dose. Details and draws from PDP1 and PDP2 are given in Appendix B.

Our objective is to vaccinate the population so as to minimize the total number of symptomatic infected individual-days over the course of the epidemic. We define a vaccination session as a tuple $(t, d, d', n)$ where $t$ is the vaccination date of the target population, $d$ is the number of doses received by individuals in the target population prior to the vaccination session, $d' > d$ is the number of doses received by each individual in the target population after the vaccination session, and $n$ is the total number of doses administered in the target population. For instance, $(2, 0, 2, 4000)$ reads: at time $t = 2$, give $d' - d = 2$

[2]Notice that the prior distributions for parameters are not necessarily independent.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

| Parameter | Description |
|---|---|
| $R_0$ | Reproduction number of unvaccinated individuals. |
| $\gamma$ | Rate of recovery. |
| $\rho_1$ | Relative infectiousness of individuals having received 1 dose of vaccine. |
| $\rho_2$ | Relative infectiousness of individuals having received 2 doses of vaccine. |
| $\sigma_1$ | Relative susceptibility of individuals having received 1 dose of vaccine. |
| $\sigma_2$ | Relative susceptibility of individuals having received 2 doses of vaccine. |
| $p_0$ | Probability of symptomatic infection of unvaccinated individuals. |
| $p_1$ | Probability of symptomatic infection of individuals having received 1 dose of vaccine. |
| $p_2$ | Probability of symptomatic infection of individuals having received 2 doses of vaccine. |

Table 1: Transmission model parameters.

doses to $n/(d' - d) = 2,000$ individuals that have not received any dose previously ($d = 0$) – a total of $n = 4000$ doses of vaccine are administered for this session.

Notice that a target population is only defined by the number $d$ of previously received doses. For simplicity, individuals to be vaccinated during the session are randomly drawn in the target population. In practice, though, vaccination campaigns could be made more efficient by using information on who is symptomatically infected and who recovered from a symptomatic infection and targeting individuals accordingly.

In general, a vaccination policy is made up of several consecutive or simultaneous vaccination sessions. Vaccination session and policies are subject to several feasibility constraints (e.g. there should be enough individuals in target populations). These constraints are given formally in Appendix C. In this article, we consider the sets of feasible alternative policies

$$\begin{aligned} A(t, n_{\max}) = \{\{(t,0,1,n_1), (t,0,2,n_2)\} \mid & \\ & n_1 \in \{k \times 2,000 \mid k \in \mathbb{N}\}, \\ & n_2 \in \{k \times 2,000 \mid k \in \mathbb{N}\}, \\ & n_1 + \frac{n_2}{2} \leq N, \\ & n_1 + n_2 \leq n_{\max}\}. \end{aligned}$$

where $t \in \mathbb{N}$ is the date at which the vaccination sessions are implemented and $n_{\max}$ is the

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

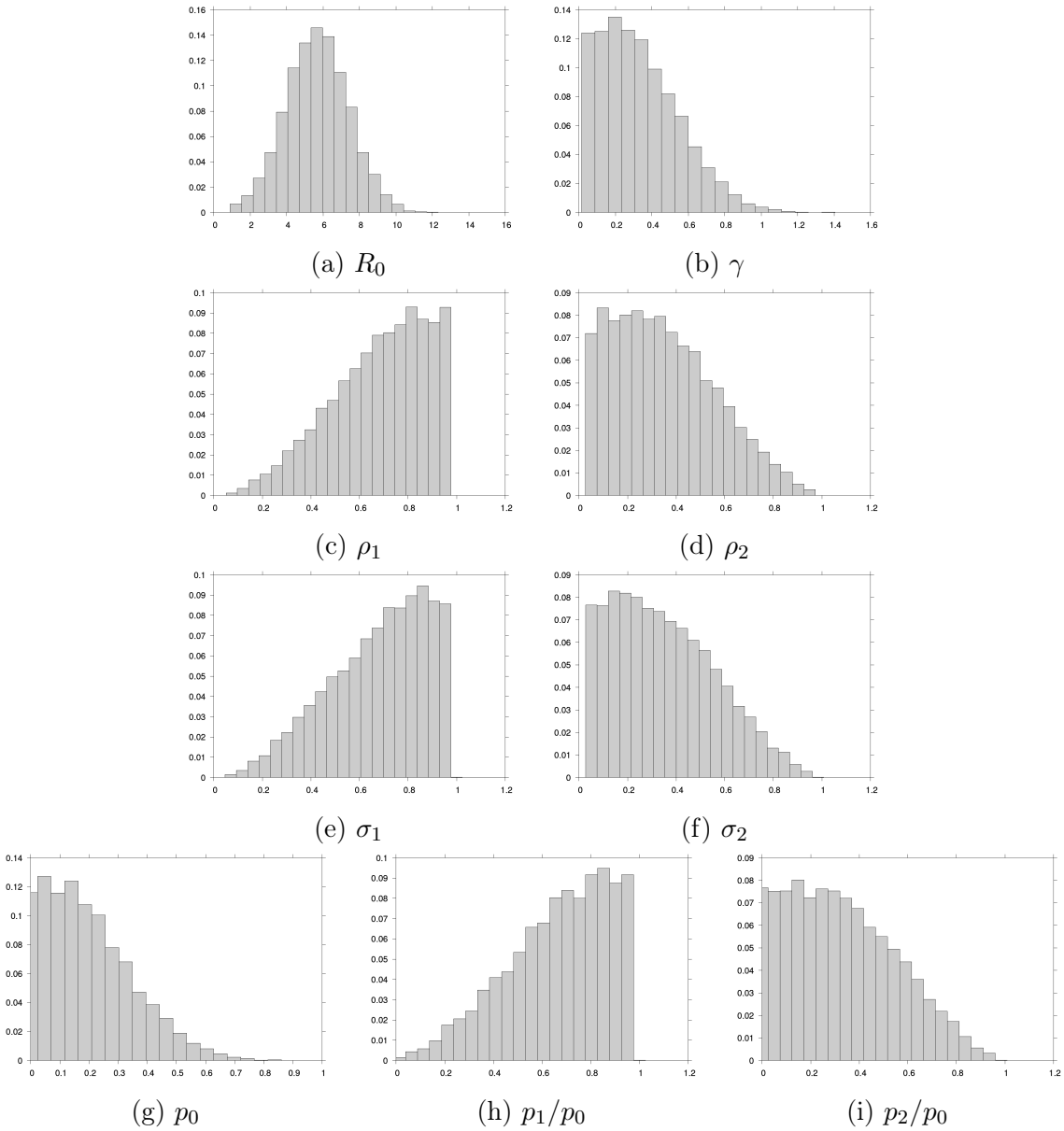



Figure 1: Parameters drawn from PDP0 (10,000 draws).

maximum number of vaccination doses administered.

Notice that we assume that vaccine distribution is done by bundles of 2,000 doses, and that $n_1$ doses are administered as single doses and $n_2$ doses as double doses (to $n_2/2$ individuals) at a single date $t$. Depending on the values of uncertain parameters (the relative efficacy of single and double dose vaccination in particular) and on feasible policies ($n_{\max}$ in particular), it is better to vaccinate more individuals with a single dose, or less individuals with a double dose.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

For each of the 10,000 draws of parameter values $\xi$ (e.g. Figure 1 for PDP0), we compute the reduction in symptomatic individual-days $V(a, \xi)$ allowed by each alternative policy in $A(t, n_{\max})$. Then, we estimate the maximum ex ante expected value and the expected ex post maximum value of epidemic control, that is the value of resolving uncertainties given the set $A(t, n_{\max})$ of feasible policies. By considering different values for $t$ and $n_{\max}$, we show the trade-off between relaxing these constraints and resolving parameter uncertainties.

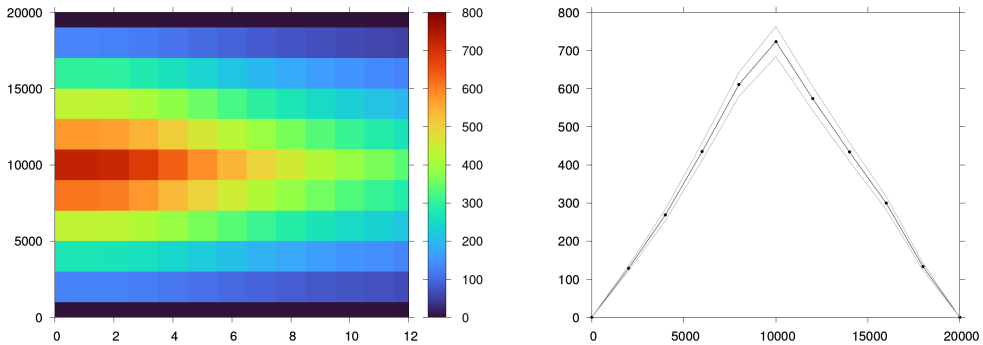
# 3 Results

Without vaccination, the expected number of symptomatic infected individual-days is 9,214.3 with 95 % confidence interval (CI) 8,992.0 – 9,436.6 for PDP0. Figure 2 shows the EVPI, that is the difference between the ex ante maximum expected value of control and the expected ex post maximum value of control, as a function of the set $A(t, n_{\max})$ of feasible policies with variable time of implementation $t$ and number of available doses $n_{\max}$. The EVPI is lower for later policy implementation dates $t$ (x-axis in Figure 2a). As time passes, the epidemic spreads and resolving uncertainties has less value. The number of available vaccine doses $n_{\max}$ is shown on the y-axis in Figure 2a and the x-axis in Figure 2b. When there is no vaccine dose available ($n_{\max} = 0$), no individual can get vaccinated with or without information, and the EVPI is zero. With 20,000 available vaccine doses, each of the $N = 10{,}000$ individuals in the population can receive two doses, so that collecting information does not bring additional value and the EVPI is zero. The EVPI is maximal for $n_{\max} = 10{,}000$ available vaccine doses. Here, 10,000 individuals can receive one dose or 5,000 individuals two doses. When less than 10,000 doses are available, the EVPI is lower because additional information cannot be fully taken advantage of due to a lack of doses. When more than 10,000 doses are available, the EVPI is lower because all individuals can receive a single dose and some of them a double dose, so the performance of policies picked based on additional information is closer to that of policies picked without resolving uncertainties. Similar results are provided for PDP1 and PDP2 in Appendix D.

We illustrated how the set of available policy options, that is constraints on policies, has an influence on the value of resolving uncertainties. In order to show the trade-off between resolving uncertainties and relaxing policy constraints, we need to compute the ex ante maximum expected value and the expected ex post maximum value separately.

Figure 3 shows the ex ante maximum expected value and the expected ex post maximum value of control as a function of policy implementation date $t$, for 10,000 available vaccine doses and PDP0 (see Appendix E for heatmaps and results for PDP1 and PDP2). For any implementation time, more information is better than less information (EVPI is

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .




(a) x-axis: $t$, y-axis: $n_{\max}$, color axis: EVPI in averted symptomatic individual-days for PDP0.

(b) $t = 0$, x-axis: $n_{\max}$, y-axis: EVPI in averted symptomatic individual-days for PDP0. Thin lines: 95% CI.

Figure 2: EVPI in averted symptomatic individual-days for PDP0 as a function of $A(t, n_{\max})$.

positive or null by definition). For a given level of information, vaccinating earlier is better than vaccinating later. This corresponds to the common intuition that early interventions prevent the subsequent spread of pathogens. However, the figure also shows that vaccinating at time $t = 0$ without additional information has the same value as vaccinating 4–5 days later with perfect information. Thus, starting from $t = 0$, collecting information is a better option than vaccinating at $t = 0$ without additional information as long as collecting information allows to resolve uncertainties in less than 4–5 days. The graph can also be read the other way round. Assume that vaccination was initially planned at $t = 4$, e.g. because the 10,000 vaccine doses will only be ready at time $t = 4$. In our scenario, it is better to spend additional resources to resolve uncertainties before $t = 4$ and vaccinate at $t = 4$ with perfect information, rather than spend them to accelerate vaccine production and bring vaccination forward (but without additional information). The same reasoning applies starting from $t = 6$. Vaccinating at time $t = 6$ with perfect information brings approximately the same value as vaccinating at time $t = 4$ without additional information. If vaccine production can be accelerated so as to vaccinate before $t = 4$, then accelerating vaccine production is the preferred option. Otherwise, it is better to resolve uncertainties and vaccinate at $t = 6$ with perfect information.

Similarly, Figure 4 shows the ex ante maximum expected value and the expected ex post maximum value of control as a function of the number of available vaccine doses $n_{\max}$ for implementation date $t = 0$ and PDP0. For any number of available vaccine doses, more information is better than (or equivalent to, for $n_{\max} = 0$ and $n_{\max} = 20{,}000$) less information. For a given level of information, vaccinating with more doses is better than

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

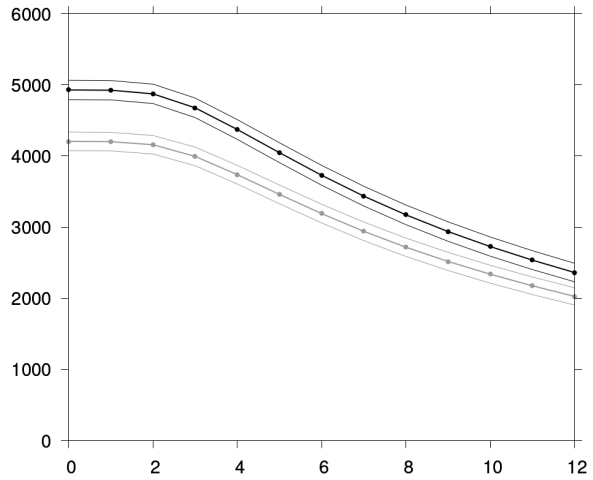



Figure 3: Ex ante maximum expected value (grey) and expected ex post maximum value (black) for PDP0 as a function of $t$ for $n_{\max} = 10{,}000$. Thin lines: 95% CI.

vaccinating with less doses. However, starting for instance from $n_{\max} = 10{,}000$ available doses and ex ante level of information, approximately the same value can be achieved by resolving uncertainties or by acquiring approximately 2,000 more doses. Thus, resolving uncertainties is a better option if it is less expensive than acquiring or producing 2,000 doses. Finally, though this is a rather far-fetched scenario, up to 2,000 doses out of 12,000 can be sold if the benefit allows to resolve uncertainties. Notice that our results remain unchanged with smaller dose bundles, see Figure App-6 in appendix.

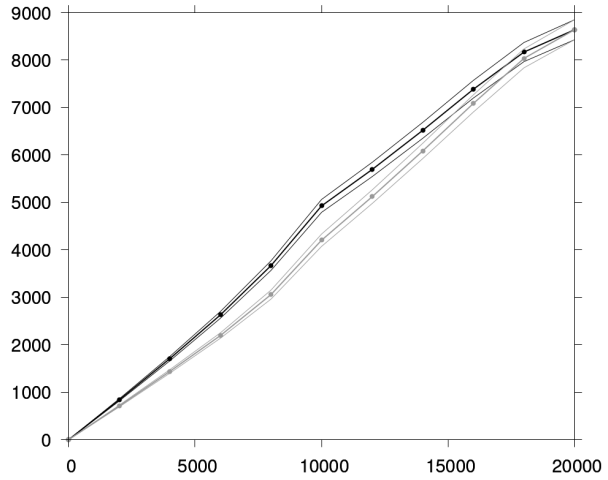



Figure 4: Ex ante maximum expected value (grey) and expected ex post maximum value (black) for PDP0 as a function of $n_{\max}$ for $t = 0$. Thin lines: 95% CI.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

# 4 Conclusion

EVPI is typically used to estimate the value of resolving uncertainties regarding an epidemic given a set of feasible control policies. In this article, we used an EVPI framework to explore the trade-off between resolving uncertainties and relaxing policy constraints. We focused on two constraints: the implementation date, and the number of available vaccine doses. We asked respectively "is it better to use resources to bring vaccination forward or to resolve uncertainties?", and "is it better to use resources to acquire more vaccine doses or to resolve uncertainties?". We first showed how the EVPI depends on policy constraints. Then, we compared the expected ex post value of control (after uncertainties are resolved) with the ex ante expected value of control (before uncertainties are resolved) for different policy constraints in order to show how relaxing constraints compares to resolving uncertainties.

Similar trade-offs have been presented in previously published studies, but most often implicitly in the context of multi-stage epidemic control. Other studies discussed such trade-offs explicitly, but only for a given type of policy constraint, not over the full range of constraints, or in very special cases. We believe that the trade-offs between relaxing policy constraints and resolving uncertainties deserved a more thorough discussion at a more abstract level.

We used a generic transmission model and a generic value of information metric to make a very general argument. We must emphasize that the EVPI is only one value of information metric among others. It assumes that uncertainties can be entirely resolved, which is unlikely to be the case in real life. Thus it only provides upper bounds on the value of collecting information. Besides, as noted by other authors [1], the time and effort required to collect information may depend on the parameters. The success of data collection may also be uncertain. To remain general, we did not look into the details of the cost of data collection and of vaccination campaigns, including the fact that producing vaccine doses may take time. These costs can be critical in practice and will depend on each specific case. Overlooking the cost details here does not make our results any less relevant since we give results in terms of upper bounds of the value of information. We also assumed that collection of information and, say, vaccination with more doses, are two mutually exclusive uses of resources. Yet vaccinating more could actually bring information about the vaccine's properties [17]. We left this scenario aside because it mostly concerns uncertainty on some of the parameters (the vaccine's properties), and because it raises the question of *how* to infer information from specific data, which is not the point of our article (we merely look at the value of information).

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Hitting fast and hard is commonly regarded as a sensible policy in the face of an epidemic. Using resources to acquire information instead, it might be argued, raises issues of ethical and political acceptability somewhat reminding of the colloquial trolley problem. Yet arguably any political or ethical constraint (or objective) can be included in a value of information analysis e.g. by choosing a performance metric and setting an objective (maximizing the expected value over uncertain parameters, minimizing the probability of adverse outcome, etc.). Thus, waving information collection aside *a priori* and without further consideration can hardly be passed off as a deliberate ethical or political stance. It rather indicates flawed decision making.

# Declarations of interest

None.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

# References

[1] Katriona Shea, Ottar N Bjørnstad, Martin Krzywinski, and Naomi Altman. Uncertainty and the management of epidemics. Nature methods, 17(9):867–869, 2020.

[2] Gianluca Baio. Statistical modeling for health economic evaluations. Annual review of statistics and its application, 5:289–309, 2018.

[3] Christopher H Jackson, Gianluca Baio, Anna Heath, Mark Strong, Nicky J Welton, and Edward CF Wilson. Value of information analysis in models to inform health policy. Annual Review of Statistics and Its Application, 9, 2021.

[4] Katriona Shea, Michael J Tildesley, Michael C Runge, Christopher J Fonnesbeck, and Matthew J Ferrari. Adaptive management and the value of information: learning via intervention in epidemiology. PLoS biology, 12(10):e1001970, 2014.

[5] Naomi V Bradbury, William JM Probert, Katriona Shea, Michael C Runge, Christopher J Fonnesbeck, Matt J Keeling, Matthew J Ferrari, and Michael J Tildesley. Quantifying the value of perfect information in emergency vaccination campaigns. PLoS computational biology, 13(2):e1005318, 2017.

[6] Nicolas Houy and Julien Flaig. Informed and uninformed empirical therapy policies. Mathematical medicine and biology: a journal of the IMA, 37(3):334–350, 2020.

[7] RN Thompson, JE Stockwin, Rolina D van Gaalen, JA Polonsky, ZN Kamvar, PA Demarsh, Elisabeth Dahlqwist, Siyang Li, Eve Miguel, Thibaut Jombart, et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. Epidemics, 29:100356, 2019.

[8] William JM Probert, Chris P Jewell, Marleen Werkman, Christopher J Fonnesbeck, Yoshitaka Goto, Michael C Runge, Satoshi Sekiguchi, Katriona Shea, Matt J Keeling, Matthew J Ferrari, et al. Real-time decision-making during emergency disease outbreaks. PLoS computational biology, 14(7):e1006202, 2018.

[9] Sibo Cheng, Christopher C Pain, Yi-Ke Guo, and Rossella Arcucci. Real-time updating of dynamic social networks for covid-19 vaccination strategies. Preprint https://arxiv. org/abs/2103, 485, 2021.

[10] Horst Behncke. Optimal control of deterministic epidemics. Optimal control applications and methods, 21(6):269–285, 2000.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

[11] B Ainseba and M Iannelli. Optimal screening in structured sir epidemics. Mathematical Modelling of Natural Phenomena, 7(3):12–27, 2012.

[12] David W Berger, Kyle F Herkenhoff, and Simon Mongey. An seir infectious disease model with testing and conditional quarantine. Technical report, National Bureau of Economic Research, 2020.

[13] Daron Acemoglu, Alireza Fallah, A Giometto, D Huttenlocher, A Ozdaglar, Francesca Parise, and Sarath Pattathil. Optimal adaptive testing for epidemic control: combining molecular and serology tests. arXiv preprint arXiv:2101.00773, 2021.

[14] Daniel B Larremore, Bryan Wilder, Evan Lester, Soraya Shehata, James M Burke, James A Hay, Milind Tambe, Michael J Mina, and Roy Parker. Test sensitivity is secondary to frequency and turnaround time for covid-19 screening. Science advances, 7(1):eabd5393, 2021.

[15] Alicia NM Kraay, Kristin N Nelson, Conan Y Zhao, David Demory, Joshua S Weitz, and Benjamin A Lopman. Modeling serological testing to inform relaxation of social distancing for covid-19 control. Nature communications, 12(1):1–10, 2021.

[16] Guanlin Li, Shashwat Shivam, Michael E Hochberg, Yorai Wardi, and Joshua S Weitz. Disease-dependent interaction policies to support health and economic outcomes during the covid-19 epidemic. Iscience, 24(7):102710, 2021.

[17] Benjamin D Atkins, Chris P Jewell, Michael C Runge, Matthew J Ferrari, Katriona Shea, William JM Probert, and Michael J Tildesley. Anticipating future learning affects current control decisions: A comparison between passive and active adaptive management in an epidemiological setting. Journal of Theoretical Biology, 506:110380, 2020.

[18] Robin N Thompson, Christopher A Gilligan, and Nik J Cunniffe. Control fast or control smart: When should invading pathogens be controlled? PLoS computational biology, 14(2):e1006014, 2018.

[19] Shou-Li Li, Matthew J Ferrari, Ottar N Bjørnstad, Michael C Runge, Christopher J Fonnesbeck, Michael J Tildesley, David Pannell, and Katriona Shea. Concurrent assessment of epidemiological and operational uncertainties for optimal outbreak control: Ebola as a case study. Proceedings of the Royal Society B, 286(1905):20190774, 2019.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

[20] Beth Woods, Claire Rothery, Sarah-Jane Anderson, Jeffrey W Eaton, Paul Revill, Timothy B Hallett, and Karl Claxton. Appraising the value of evidence generation activities: an hiv modelling study. BMJ global health, 3(6):e000488, 2018.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

# Appendix

## A Compartmental model

Let $S_i$ and $I_i$ denote the number of susceptible and infectious individuals who received $i \in \{0, 1, 2\}$ doses of vaccine, and $R$ the number of recovered individuals. We assume a closed population of $N$ individuals.

Individuals get infected with contact rate $\beta$ and they recover at rate $\gamma$. The basic reproduction number is $R_0 = \beta/\gamma$. Individuals who received $i \in \{1, 2\}$ vaccine doses are less susceptible by a factor $\sigma_i$ and less infectious by a factor $\rho_i$ compared to unvaccinated individuals. We assume that vaccination is instantaneous and that there is no waning of vaccine protection. The epidemic dynamics is described by Equations (1)–(5).

$$\frac{dS_0}{dt} = -\lambda S_0 \tag{1}$$

$$\frac{dI_0}{dt} = \lambda S_0 - \gamma I_0 \tag{2}$$

$$\frac{dS_i}{dt} = -\sigma_i \lambda S_i \tag{3}$$

$$\frac{dI_i}{dt} = \sigma_i \lambda S_i - \gamma I_i \tag{4}$$

$$\frac{dR}{dt} = \gamma \sum_{j=0}^{2} I_j \tag{5}$$

where

$$\lambda = \beta(I_0 + \rho_1 I_1 + \rho_2 I_2)/N. \tag{6}$$

Let $p_i$ the probability of symptomatic infection for infected individuals who received $i$ doses of vaccine. The incidence rates of symptomatic and asymptomatic cases among unvaccinated individuals are $\mathcal{I}_0^{\mathrm{s}} = p_0 \lambda S_0$ and $\mathcal{I}_0^{\mathrm{as}} = (1 - p_0)\lambda S_0$ respectively. Among individuals who received $i \in \{1, 2\}$ doses of vaccine, incidence rates are $\mathcal{I}_i^{\mathrm{s}} = p_i \sigma_i \lambda S_i$ and $\mathcal{I}_i^{\mathrm{as}} = (1 - p_i)\sigma_i \lambda S_i$.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

| Parameter | Distribution | Valid values |
|---|---|---|
| $R_0$ | $\mathcal{N}(6, 3)$ | $[1.2, \infty[$ |
| $\gamma$ | $\mathcal{N}(0.2, 0.1)$ | $[0.05, \infty[$ |
| $\rho_1$ | $\mathcal{N}(0.75, 0.25)$ | $[\rho_2, 1]$ |
| $\rho_2$ | $\mathcal{N}(0.75, 0.25)$ | $[0.05, \rho_1]$ |
| $\sigma_1$ | $\mathcal{N}(0.75, 0.25)$ | $[\sigma_2, 1]$ |
| $\sigma_2$ | $\mathcal{N}(0.75, 0.25)$ | $[0.05, \sigma_1]$ |
| $p_0$ | $\mathcal{N}(0.1, 0.05)$ | $[0, 1]$ |
| $p_1/p_0$ | $\mathcal{N}(0.75, 0.25)$ | $[p_2/p_0, 1]$ |
| $p_2/p_0$ | $\mathcal{N}(0.4, 0.25)$ | $[0, p_1/p_0]$ |

Table App-1: Prior distributions for parameters 0 (PDP0). $\mathcal{N}(\mu, \sigma^2)$: normal distribution with mean $\mu$ and variance $\sigma^2$.

| Parameter | Distribution |
|---|---|
| $R_0$ | $\mathcal{U}(1.2, 10)$ |
| $\gamma$ | $\mathcal{U}(0.05, 0.5)$ |
| $\rho_1$ | $\mathcal{U}(\rho_2, 1)$ |
| $\rho_2$ | $\mathcal{U}(0.05, \rho_1)$ |
| $\sigma_1$ | $\mathcal{U}(\sigma_2, 1)$ |
| $\sigma_2$ | $\mathcal{U}(0.05, \sigma_1)$ |
| $p_0$ | $\mathcal{U}(0, 1)$ |
| $p_1/p_0$ | $\mathcal{U}(p_2/p_0, 1)$ |
| $p_2/p_0$ | $\mathcal{U}(0, p_1/p_0)$ |

Table App-2: Prior distributions for parameters 1 (PDP1). $\mathcal{U}(a, b)$: uniform distribution on $[a, b]$.

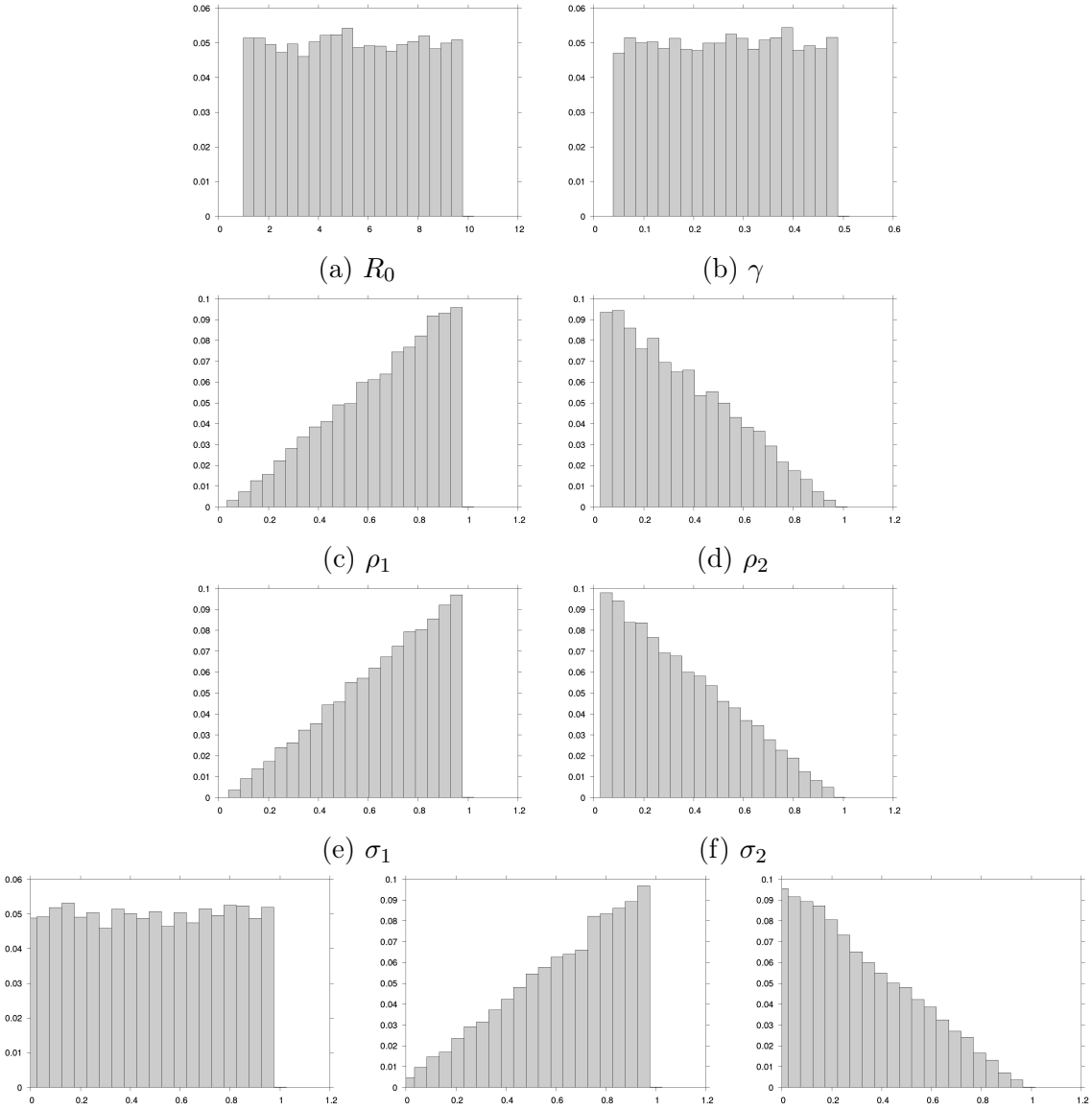
# B Prior distributions for parameters

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

| **Parameter** | **Distribution** | **Valid values** |
|---|---|---|
| $R_0$ | $\mathcal{N}(6, 3)$ | $[1.2, \infty[$ |
| $\gamma$ | $\mathcal{N}(0.2, 0.1)$ | $[0.05, \infty[$ |
| $\rho_1$ | $\mathcal{N}(0.9, 0.25)$ | $[\rho_2, 1]$ |
| $\rho_2$ | $\mathcal{N}(0.2, 0.25)$ | $[0.05, \rho_1]$ |
| $\sigma_1$ | $\mathcal{N}(0.9, 0.25)$ | $[\sigma_2, 1]$ |
| $\sigma_2$ | $\mathcal{N}(0.2, 0.25)$ | $[0.05, \sigma_1]$ |
| $p_0$ | $\mathcal{N}(0.1, 0.05)$ | $[0, 1]$ |
| $p_1/p_0$ | $\mathcal{N}(0.9, 0.25)$ | $[p_2/p_0, 1]$ |
| $p_2/p_0$ | $\mathcal{N}(0.2, 0.25)$ | $[0, p_1/p_0]$ |

Table App-3: Prior distributions for parameters 2 (PDP2). $\mathcal{N}(\mu, \sigma^2)$: normal distribution with mean $\mu$ and variance $\sigma^2$.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .



(a) $R_0$ (b) $\gamma$ (c) $\rho_1$ (d) $\rho_2$ (e) $\sigma_1$ (f) $\sigma_2$ (g) $p_0$ (h) $p_1/p_0$ (i) $p_2/p_0$

Figure App-1: Parameters drawn from PDP1 (10,000 draws).

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

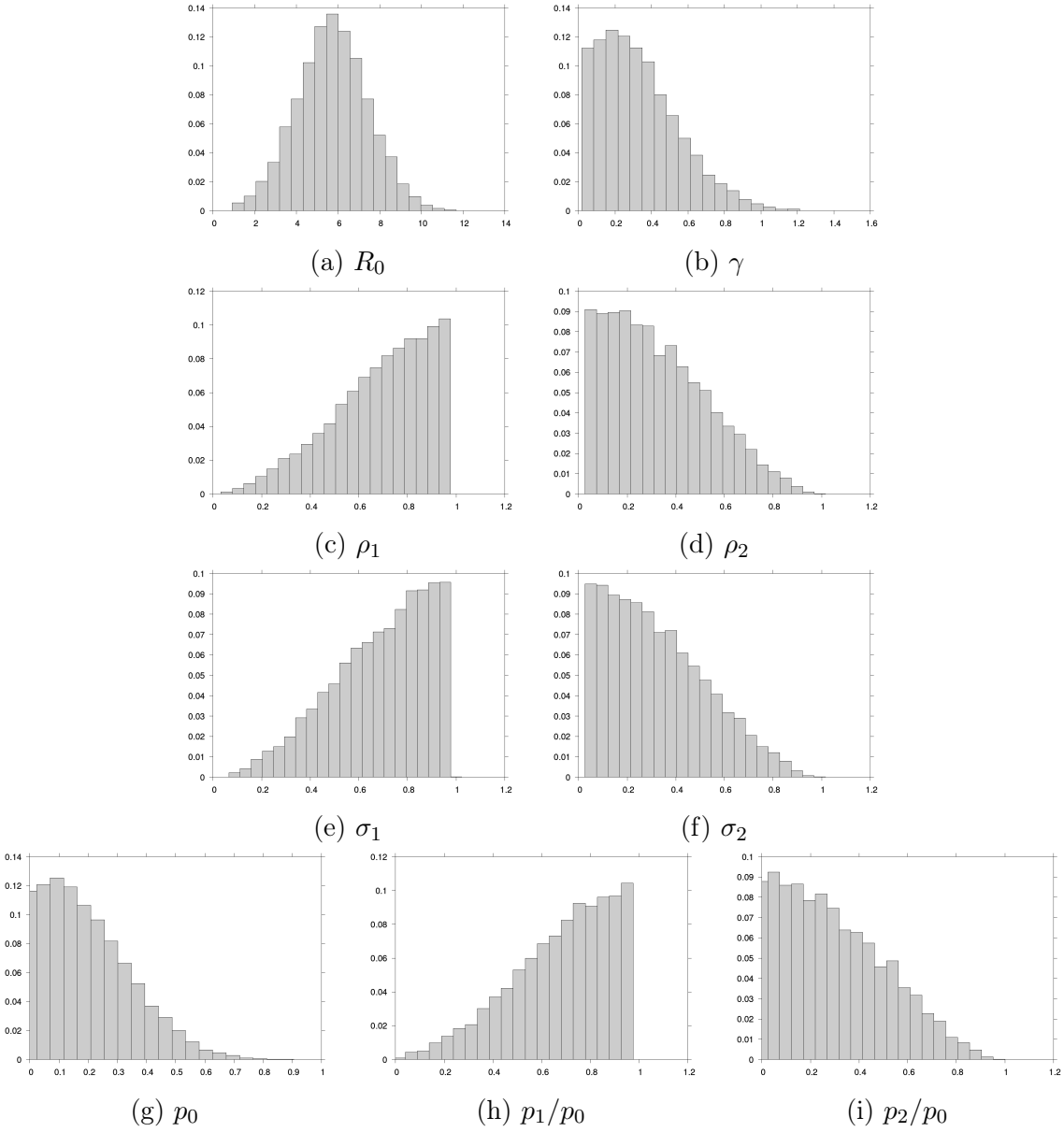



Figure App-2: Parameters drawn from PDP2 (10,000 draws).

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

# C Feasibility of vaccination policies

Let $a = \{(t_0, d_0, d'_0, n_0), \dots (t_{K-1}, d_{K-1}, d'_{K-1}, n_{K-1})\}$ a vaccination policy made up of $K$ vaccination sessions. For the sake simplicity, we restrict to the case where time is discontinuous – *i.e.* $\forall s_i = (t_i, d_i, d'_i, n_i) \in a, t_i \in \mathbb{N}$ – and where individuals can received up to 2 vaccine doses – *i.e.* $\forall s_i = (t_i, d_i, d'_i, n_i) \in a, (d_i, d'_i) \in \{(0,1), (1,2), (0,2)\}$.

For the sake of simplicity and with a slight abuse of notation, we say that $(t, d, d') \notin a$ if $\forall n \in \mathbb{N}, (t, d, d', n) \notin a$. We define $\mathcal{A}(t, d, d') = \sum_{s_i=(t_i,d_i,d'_i,n_i)\in a} I_{t_i=t} I_{d_i=d} I_{d'_i=d'} n_i$, where $I$ is the indicator function. Hence, $\mathcal{A}(t, d, d')$ is the number of doses administered at time $t$ to individuals who previously received $d$ doses and currently receive $d' - d$ doses.

We call $\mathcal{N}(d, t)$ with $d \in \{0, 1, 2\}$ and $t \in \{-1\} \cup \mathbb{N}$, the number of individuals who received $d$ doses before date $t$. Formally, we define $\mathcal{N}$ recursively with:

$$\begin{array}{rrcl}
 & \mathcal{N}(0,-1) & = & N, \\
 & \mathcal{N}(1,-1) & = & 0, \\
 & \mathcal{N}(2,-1) & = & 0, \\
\forall t \in \mathbb{N}, & \mathcal{N}(0,t) & = & \mathcal{N}(0,t-1) - \mathcal{A}(t,0,1) - \mathcal{A}(t,0,2)/2, \\
\forall t \in \mathbb{N}, & \mathcal{N}(1,t) & = & \mathcal{N}(1,t-1) + \mathcal{A}(t,0,1) - \mathcal{A}(t,1,2), \\
\forall t \in \mathbb{N}, & \mathcal{N}(2,t) & = & \mathcal{N}(2,t-1) + \mathcal{A}(t,1,2) + \mathcal{A}(t,1,2)/2.
\end{array}$$

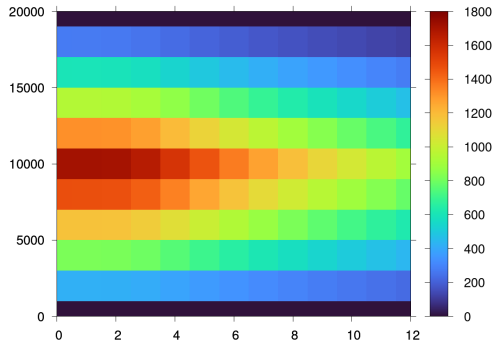
We say that $a$ is feasible if the following constraint is met: $\forall t \in \mathbb{N}, \mathcal{N}(0,t) \geq 0, \mathcal{N}(1,t) \geq 0, \mathcal{N}(2,t) \geq 0$. Notice that by definition $\forall t \in \mathbb{N}, \mathcal{N}(0,t) + \mathcal{N}(1,t) + \mathcal{N}(2,t) = N$ so that it is unnecessary to check that each vaccination state subpopulation is inferior to $N$.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

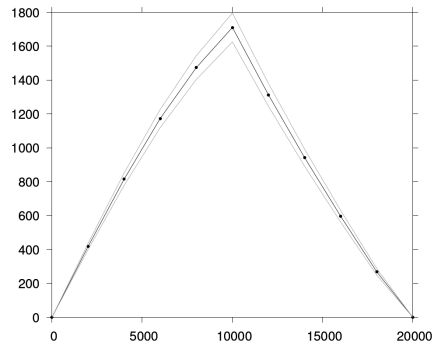
# D Additional figures: EVPI

The expected total infected individual-days is 24,051.9 (95 % CI : 23,569.0 – 24,534.7) with PDP1. The expected total infected individual-days is 9,342.8 (95 % CI : 9,117.0 – 9,568.6) with PDP2.

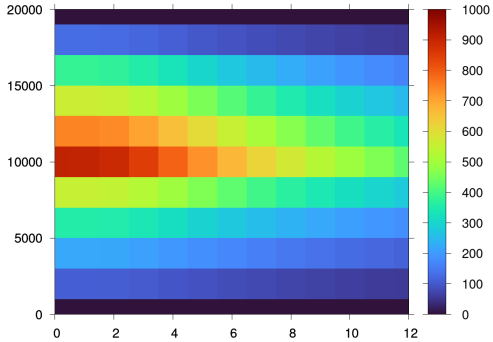
Figure App-3 shows the EVPI as a function of feasible policies $A(t, n_{\max})$ for PDP1 and PDP2.



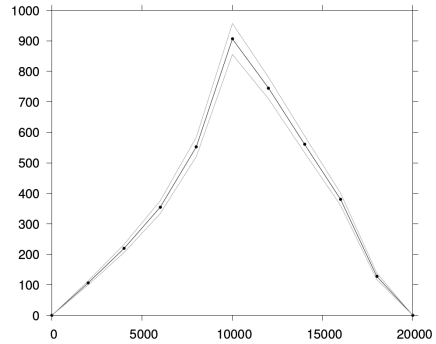

(a) PDP1, x-axis: $t$, y-axis: $n_{\max}$.




(b) PDP1, $t = 0$, x-axis: $n_{\max}$. Thin lines: 95% CI.




(c) PDP2, x-axis: $t$, y-axis: $n_{\max}$.




(d) PDP2, $t = 0$, x-axis: $n_{\max}$. Thin lines: 95% CI.

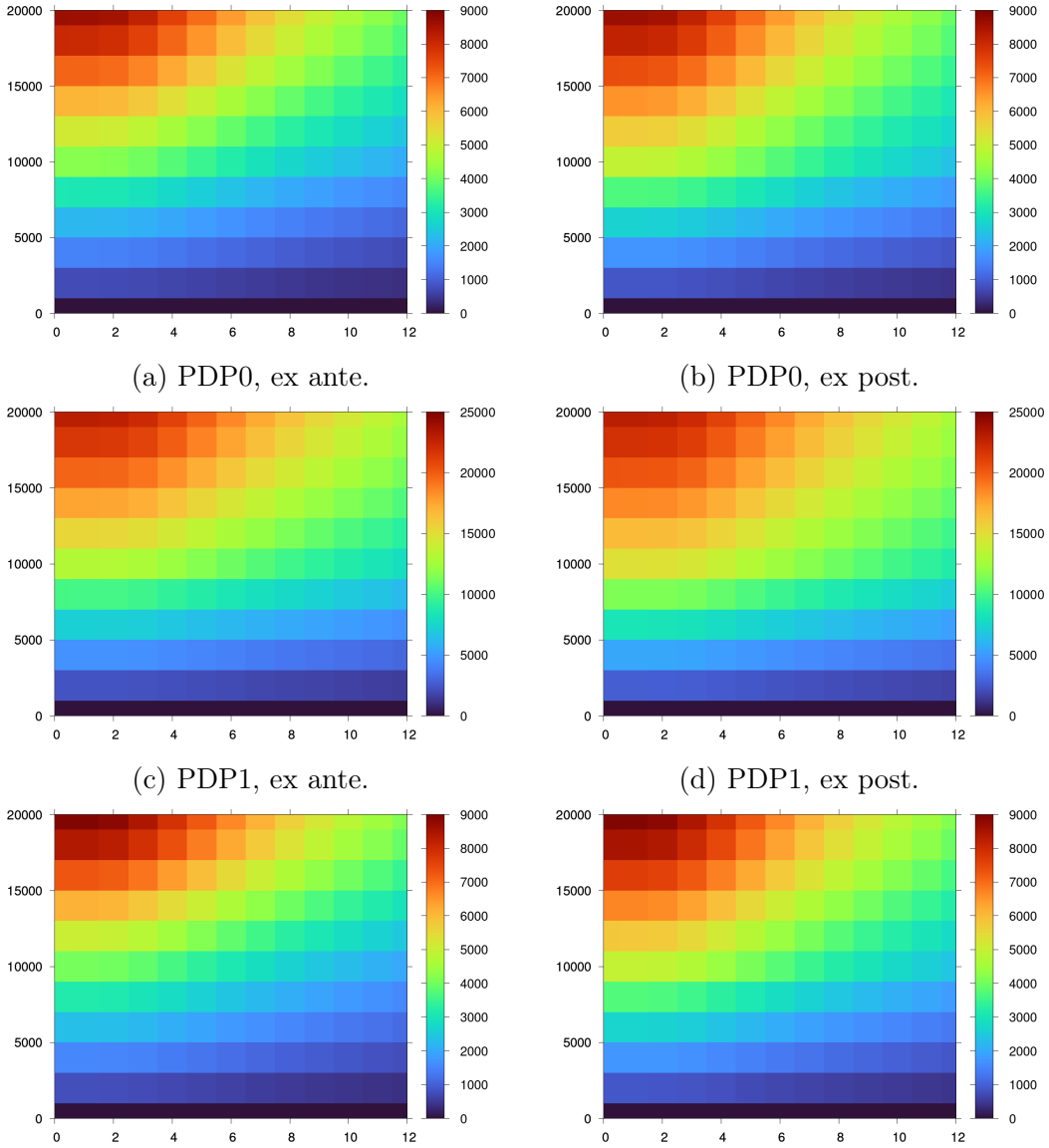
Figure App-3: EVPI in averted symptomatic individual-days for PDP1 and PDP2 as a function of $A(t, n_{\max})$.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

# E Additional figures: ex ante maximum expected value v. expected ex post maximum value

Figure App-4 shows the ex ante maximum expected value $\max_{a \in A} E_\xi \left[ V(a, \xi) \right]$ and the expected ex post maximum value $E_\xi \left[ \max_{a \in A} V(a, \xi) \right]$ as a function of $A(t, n_{\max})$.

medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .




(a) PDP0, ex ante.

(b) PDP0, ex post.

(c) PDP1, ex ante.

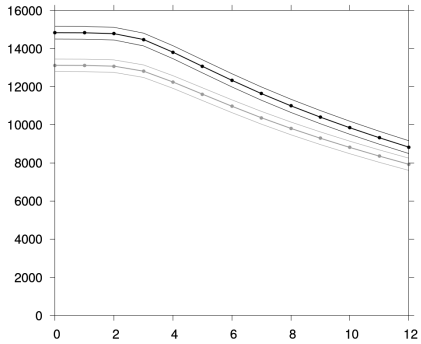
(d) PDP1, ex post.

(e) PDP2, ex ante.

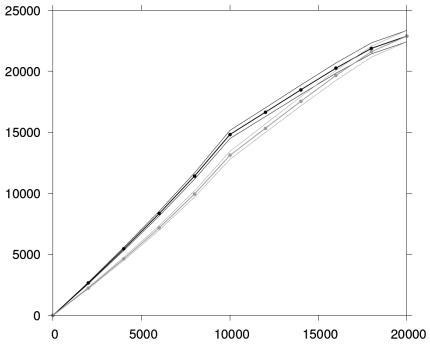
(f) PDP2, ex post.

Figure App-4: Ex ante maximum expected value and expected ex post maximum value for PDP0, PDP1, and PDP2 as a function of $A(t, n_{\max})$. x-label: policy implementation date $t$. y-label: $n_{\max}$.

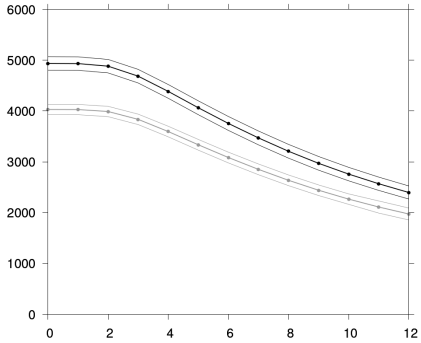
medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .



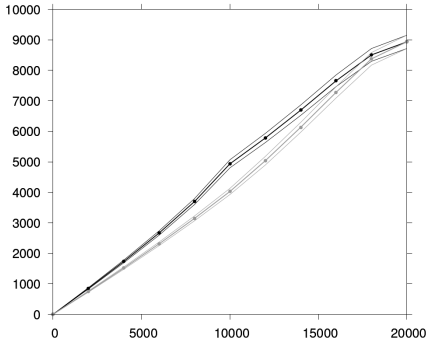

(a) PDP1, $n_{\max} = 10{,}000$, x-axis: $t$.




(b) PDP1, $t = 0$, x-axis: $n_{\max}$.



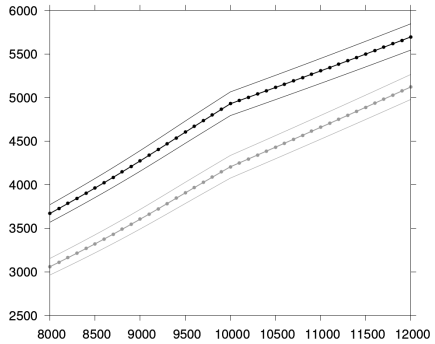

(c) PDP2, $n_{\max} = 10{,}000$, x-axis: $t$.



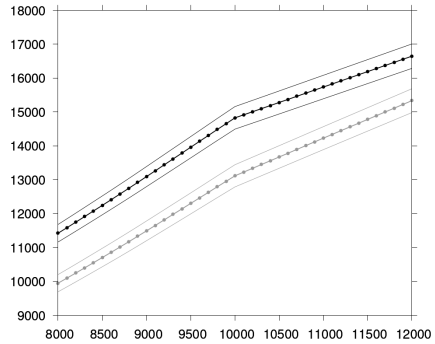

(d) PDP2, $t = 0$, x-axis: $n_{\max}$.

Figure App-5: Ex ante maximum expected value (grey) and expected ex post maximum value (black) for PDP1 and PDP2, as a function of $t$ for $n_{\max} = 10{,}000$ and as a function of $n_{\max}$ for $t = 0$. Thin lines: 95% CI.

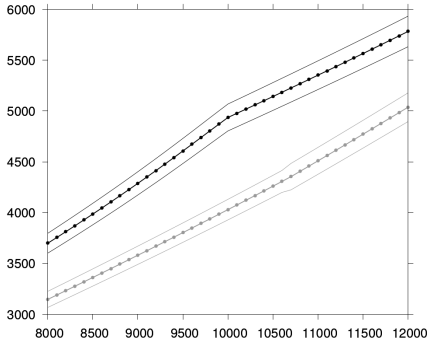
medRxiv preprint doi: https://doi.org/10.1101/2022.03.28.22273039; this version posted March 30, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .




(a) PDP0.




(b) PDP1.




(c) PDP2.

Figure App-6: Ex ante maximum expected value (grey) and expected ex post maximum value (black) for PDP0, PDP1, and PDP2 as a function of $n_{\max}$ for $t = 0$. Thin lines: 95% CI.