Expanding Capacity for Vaccines Against Covid-19 and Future Pandemics: A Review of Economic Issues

Susan Athey, Juan Camilo Castillo, Esha Chaudhuri, Michael Kremer, Alexandre Simoes Gomes, and Christopher M. Snyder

JULY 2022
EXPANDING CAPACITY FOR VACCINES AGAINST COVID-19
AND FUTURE PANDEMICS:
A REVIEW OF ECONOMIC ISSUES

Susan Athey
Juan Camilo Castillo
Esha Chaudhuri
Michael Kremer
Alexandre Simoes Gomes
Christopher Snyder

July 2022

This paper synthesizes a series of articles written by members of the Accelerating Health Technologies (AHT) team (www.acceleratinght.org), which besides the authors includes Amrita Ahuja, Arthur Baker, Owen Barder, Eric Budish, Tasneem Chipty, Rachel Glennerster, Scott Kominers, Greg Larson, Jean Lee, Jonathan Levin, Anup Malani, Canice Prendergast, Alex Tabarrok, Brandon Tan, Duc Tran, and Witold Więcek. AHT’s work was supported in part by funding from the Wellspring Philanthropic Fund (grant no. 15104) and Open Philanthropy and aided by advice and effort from Whit Athey, Lisa Danzig, Dimitrios Gouglas, Kendall Hoyt, and Jessica Picket. The paper was substantially improved following comments from editor Alex Tabarrok, discussant Christopher Adam, and an anonymous referee. The authors are grateful for helpful comments from Witold Więcek and participants in the “Economics of Pandemic Vaccination” seminar sponsored by the Oxford Review of Economic Policy and the “Medical Supply Chain Security and Governance” conference sponsored by the Global Health Initiative of Dartmouth's Dickey Center. The authors are grateful to Brandon Chen for excellent research assistance.

© 2022 by Susan Athey, Juan Camilo Castillo, Esha Chaudhuri, Michael Kremer, Alexandre Simoes Gomes, and Christopher Snyder. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.
Expanding Capacity for Vaccines Against Covid-19 and Future Pandemics: A Review of Economic Issues
Susan Athey, Juan Camilo Castillo, Esha Chaudhuri, Michael Kremer, Alexandre Simoes Gomes, and Christopher Snyder
July 2022
JEL No. D47,H44,I15,I18,O31

ABSTRACT

We review economic arguments for using public policy to accelerate vaccine supply during a pandemic. Rapidly vaccinating a large share of the global population helps avoid economic, mortality, and social losses, which in the case of Covid-19 mounted into trillions of dollars. However, pharmaceutical firms are unlikely to have private incentives to invest in vaccine capacity at the socially optimal scale and speed. The socially optimal level of public spending may cause some sticker shock but—as epitomized by the tagline “spending billions to save trillions”—is eclipsed by the benefits and can be restrained with the help of careful policy design and advance preparations. Capacity is so valuable during a pandemic that fractional dosing and other measures to stretch available capacity should be explored.

Susan Athey
Graduate School of Business
Stanford University
655 Knight Way
Stanford, CA 94305
and NBER
athey@stanford.edu

Michael Kremer
University of Chicago
Kenneth C. Griffin Department of Economics
1126 E. 59th St.
Chicago, IL 60637
and NBER
kremermr@uchicago.edu

Juan Camilo Castillo
Department of Economics
University of Pennsylvania
jccast@upenn.edu

Alexandre Simoes Gomes
Development Innovation Lab
University of Chicago
asimoes@uchicago.edu

Esha Chaudhuri
Harvard University
echaudhuri@college.harvard.edu

Christopher Snyder
Department of Economics
Dartmouth College
301 Rockefeller Hall
Hanover, NH 03755
and NBER
chris.snyder@dartmouth.edu
1. Introduction

Each month, the Covid-19 pandemic caused tens of millions of hospitalizations (Our World in Data 2022b), hundreds of thousands of deaths (Our World in Data 2022b), and hundreds of billions of dollars in lost economic output (International Monetary Fund 2020). The damage was much greater accounting for human capital losses from school closures and job separations and indirect health harms in the form of impaired mental health and worsened comorbidities (Cutlers and Summers 2020, Hanushek and Woessman 2020).

Vaccines proved to be a vital tool in containing the worst damage from the pandemic. Though the speed of vaccine development and scale-up of manufacturing capacity were unprecedented, still, public and private investment in Covid-19 vaccines were likely suboptimal, considering the enormous social benefits associated with accelerating vaccination. Vaccine supplies were extremely tight early in the pandemic even in high-income countries (HICs), for example leading the European Union to sue Astra-Zeneca for delaying shipments there while fulfilling orders in its home country, the United Kingdom (Guarascio and Smouth 2021). Coverage rates in low-income countries (LICs) and lower-middle income countries (LMICs) continued to lag those in HICs well into the pandemic, with many of those countries failing to vaccinate even 5% of their populations by the end of 2021.¹ Surveys finding a higher willingness to take Covid-19 vaccines in LICs and LMICs compared to some HICs (Solís Arce et al. 2021) suggest that scarce supply rather than scarce demand was the cause of low initial uptake there. Although vaccine supply is not currently the main constraint to vaccination, the emergence of new variants (with potential reallocation of manufacturing capacity to vaccines developed for those variants), supply shocks, and the rising demand for boosters in higher-income countries may further delay vaccine supplies to these countries.

The shortage of vaccine capacity can be traced to the gap between social and commercial incentives for vaccine capacity. Political and social pressure constrain the prices that can be charged for vaccines, preventing them from coming anywhere close to their full social value. The prices observed in vaccine deals to date, ranging from $4 to $60 per course, fall far short of estimates of the social value of capacity from Castillo et al. (2021), around $6,200 per annual course. Even higher

¹As of June 2022, 54% of the population in LMICs have received a complete initial vaccine protocol but only 13% in LICs have (Our World in Data 2022c). A complete protocol is two doses for most Covid-19 vaccines; a person who has completed the initial protocol has not necessarily received boosters.
prices would not provide the full social incentives to accelerate vaccine supply if the price is the same whether courses are delivered in a month or a year.

This paper surveys economic arguments for using public policy to accelerate vaccine supply during a pandemic, with particular focus on a series of articles by the Accelerating Health Technologies team, of which this paper’s authors are members. The team’s work naturally focused on mitigating the Covid-19 pandemic peaking at that time, but the lessons from that work are relevant for any future pandemic. The arguments made here that preparations in advance of a pandemic can foster more international cooperation and involve lower expense than investments undertaken during the heat of a crisis, suggest that there are high potential returns to investing in preparations for the next pandemic—which may come sooner rather than later, perhaps in the form of the next Covid-19 wave.

We provide back-of-the-envelope calculations demonstrating that the social value of accelerating vaccine supply by incentivizing investment in capacity at risk—in parallel with clinical trials rather than delaying until after vaccine approval—can be well in the trillions of dollars. Similar calculations can be used to demonstrate trillions of dollars of social value from expanding vaccine capacity. The reason more vaccine capacity is so valuable is that scale equals speed: if scarce capacity is the rate-limiting factor, doubling capacity cuts the time to supply a given quantity of vaccines in half, accelerating vaccination campaigns and mitigating cumulating losses that much faster. The back-of-the-envelope calculations serve a pedagogical purpose but in fact are not far off the more detailed estimates from Castillo et al. (2021).

Accelerating vaccination with more capacity improves not just social welfare but equity. If, for example, there is a year-long queue, doubling capacity and halving time until the full population is vaccinated would reduce the wait time of someone one month into the queue by just two weeks. However, it would reduce the wait time of someone at the back of the queue by six months.

Having made the case that vaccine capacity is enormously valuable, eclipsing commercial incentives for scale and speed, the paper turns to the question of how to determine the optimal size of the vaccine portfolio supported by public expenditures. We focus on the approach due to Ahuja et al. (2021), which combines estimates of probabilities of success allowed to be correlated across different vaccine technologies (e.g., viral vector versus mRNA) along with estimates of capacity costs and avoided harms from vaccination to compute the number of candidates an individual country or
coalition should invest in and how much at-risk capacity to install for each. According to this approach, even some of the most aggressive country programs fell short of the optimal investment. For example, Operation Warp Speed in the United States spent $18 billion on six vaccine candidates, whereas according to Ahuja et al. (2021) over three times that amount should have been spent on 27 candidates.

Given the large expenditures involved in the policy that our model suggests is optimal, measures should be taken to ensure the expenditures are spent efficiently, not wasted. Section 6 discusses some of these design features, including combining some of both “push” funding (contributing toward capacity costs) and “pull” funding (paying a bonus price for approved vaccines produced on an expedited schedule) and contracting on capacity rather than doses.

Policies to expanded capacity would mitigate but not eliminate scarcity. Section 7 discusses measures to stretch scarce capacity, including a “first doses first” strategy, delaying second doses of a two-dose sequence to speed wider coverage of first doses, and fractional dosing, say doubling effective capacity by using half the active ingredient in each dose. Such policies may be able to expand coverage with little sacrifice of efficacy. If there are doubts about the tradeoff, there is option value in testing dose-stretching policies or rolling them out on an experimental basis. If anticipated benefits do not materialize, the program can be quickly abandoned and individuals receiving the stretched doses protected with full doses.

The concluding section summarizes implications for the Covid-19 pandemic and lessons for future pandemics. We speculate on reasons for the gap between actual levels of country procurement and the levels that the Accelerating Health Technologies team’s models suggested were optimal.

Our exclusive focus in this paper, vaccine capacity, is but one piece of a larger pandemic puzzle. Vaccinating an individual requires the dose to be manufactured, delivered, and administered to the patient, requiring investments not just in production lines but in input supply chains, transportation networks, clinics, and consumer acceptance. Vaccines are but one defense against pandemics; others include disease surveillance, diagnostics, testing, drug therapies, personal protective equipment, and social distancing. We focus on vaccines because of their long record of high benefit-cost ratios for other diseases. Their administration may require less interference with normal life than some other interventions, potentially reducing the economic costs of the pandemic. We further narrow the focus to production capacity for vaccines for several reasons. First, capacity decisions are made by
pharmaceutical manufacturers, whose commercial interests may require inducements to align with social goals. Second, production capacity may involve more economic risk and lead time than other investments involved in a vaccination campaign, so may be the rate-limiting step, at least early in a pandemic. We maintain a broad notion of production capacity including not only facility space but also sufficient staff to ramp up production at the facility and maintain quality control.

2. Valuing Capacity Investment

To help understand the benefits of expanding and accelerating vaccine capacity, consider the numerical example illustrated in Figure 1. While it is somewhat of a toy example, with assumptions adopted mainly for pedagogical convenience, we will see that its results are not too far from more detailed estimates of the value of capacity from the literature. Adopting the conservative assumption from the introduction that the pandemic causes $1 trillion of harm each month from all sources (economic, morbidity, and mortality) and supposing the pandemic lasts two years, the gray rectangle then represents the total global harm from the pandemic.

Imagine a vaccine is developed that can end the pandemic once 70% of the population is covered, what many epidemiologists early in the Covid-19 pandemic took to be the critical threshold for vaccine coverage needed to achieve herd immunity. For a variety of reasons, including the emergence of more transmissible Covid-19 strains against which vaccines proved to be less effective in preventing transmission than serious illness, epidemiologists moved away from herd immunity targets toward achieving as wide coverage as possible (D’Souza and Dowdy 2021), but this toy example adopts the earlier perspective when capacity investments were being made. Suppose the vaccine receives regulatory approval after succeeding in clinical trials at date 0, and it takes three further months for the manufacturer to install capacity for a global rollout.

We begin by assuming for simplicity that the vaccine reduces monthly global harm linearly, in proportion to its coverage up to the 70% target, as might be the case, for example, if the vaccine is randomly distributed rather than prioritizing the most vulnerable or biggest spreaders. The available

---

2See the survey in Lipsitch (2020). Herd immunity is achieved when enough individuals have been removed from the susceptible population that the infections wane rather than grow. Lipsitch (2020) provides the textbook formula for critical vaccine coverage \((1 - 1/R_0)/x\), where \(R_0\) is the disease’s basic reproductive number, a standard measure of disease transmissibility, and \(x\) is the vaccine’s efficacy in preventing transmission. Substituting reasonable estimates from that time of \(R_0 = 2.2\) and \(x = 0.75\) (see references in Lipsitch 2020) yields a critical vaccine coverage of 70%.
capacity can be used to vaccinate a cumulating number of people, cutting wedge A out of pandemic harm in Figure 1. The figure implicitly assumes that capacity is sufficient to vaccinate 70% of the population by the end of month 24, 21 months after vaccine approval.

The world may be able to do better than a linear reduction in monthly harm by better targeting, perhaps giving priority to medical personnel, the aged, or other vulnerable demographic groups. According to one estimate (Bubar et al. 2021), 80% of Covid-19 mortality reduction would come from vaccinating the first 20% of a population. Additional measures can be taken to concentrate supplies where infection rates are rising fastest or where economic losses are most severe. In the figure, better targeting avoids additional harm B. The convex border of B reflects the fact that initial courses going to the highest priorities relieve more harm than later courses going to lower priorities. We model the curve as a simple quadratic.

So far, the example has assumed that the manufacturer waits until the vaccine succeeds in trials before investing in global capacity. If three months are required to build capacity, the global rollout cannot begin in earnest until month 3. Instead of waiting for the outcome of clinical trials, the manufacturer could engage in at-risk capacity investment in parallel with clinical trials rather than sequentially. Supposing trials take three months and capacity investment can start soon after trials start, at-risk investment accelerates the availability of capacity by three months, speeding vaccinations by three months, and avoiding three months of additional harm for those vaccinated. In the figure, at-risk investment shows up as a parallel shift left in the curve delineating avoided harm, reducing global harm by the area of C. The unusual sabre shape of C belies the fact that its area is simply the $3 trillion gain obtained from accelerating the end of a pandemic costing the world $1 trillion per month by three months.

Region D shows the additional harm avoided if at-risk investment is undertaken at double the scale. Capacity equals speed: instead of taking 21 months to reach 70% global coverage, doubling capacity cuts the time in half to $21 \div 2 = 10.5$ months. While doubling capacity avoids substantial additional harm, it does not double the harm avoided. The main reason is that the example assumes enough capacity before doubling that vaccinations of the target population can be completed before the end of the pandemic. Doubling capacity does not increase the proportion of the population ultimately covered, just the speed that that coverage is attained.

There are enough concrete numbers behind the example that the value of various changes to
capacity policy can be estimated by simply calculating the area of the various regions in Figure 1. The first set of rows in Table 1 report these calculations. Investing in baseline capacity not at risk avoids $14.0 trillion of global harm (area of $A$ and $B$). Investing this capacity at risk saves an additional $3.0 trillion of harm (area of $C$) and doubling this at-risk capacity an additional $3.5 trillion of harm (area of $D$).

While the assumptions behind the numerical example were chosen for pedagogical convenience, the results are not far from more detailed estimates from Castillo et al. (2021), reported in the lower rows in Table 1. From their study of announced bilateral deals with vaccine manufacturers, Castillo et al. (2021) calculated a range for global capacity, centered on 3.0 billion courses annually, that would be consistent with those production plans. The authors projected the rollout of the vaccines produced with this capacity in proportion to the deals signed with high-, middle-, and low-income countries. They took account of nonlinear benefits arising from the rollout across different demographic groups within countries reflecting epidemiological evidence on relative harm avoided across groups, with some consideration for epidemiological externalities across groups. Despite the additional detail in Castillo et al. (2021) compared to the numerical example, there is a limit to how different their results can be given that the two analyses involve similar underlying capacities and timeframes. The detailed methodology in Castillo et al. (2021) mainly allows the convex curve bounding $B$ to be estimated more precisely than simply positing a quadratic. Still, the quadratic will be close to whatever curve bounding $B$ is ultimately estimated.

According to the estimates in Castillo et al. (2021), the baseline level of at-risk capacity installed globally saved $18.8 trillion of global harm, or about $6,200 per course of annual capacity.\(^3\) Investing in that capacity at risk rather than waiting three months accounts for an additional $18.8−15.8=3.0$ trillion dollars of avoided global harm, or $1,000$ additional benefit per course of annual capacity thus accelerated. Had at-risk capacity been doubled to 6.0 billion courses annually, global benefit would have been higher by $21.4−18.8=2.6$ trillion dollars, or about $870$ per annual course of that additional capacity.

\(^3\)These estimates are taken from Table S2 from the online supplement to Castillo et al. (2021), not the headline estimates from Table 1 of that paper of a global benefit of $17.4$ trillion, $5,800$ per course of baseline capacity. We cite the Table S2 estimate since it assumes at-risk capacity all comes online in the same month, consistent with the numerical example. The Table 1 estimate assumes capacity comes online in two tranches, which is inconsistent with the numerical example but a realistic depiction of how capacity actually ramped up at the time. The alternative estimates are fairly close in any event.
A lesson for economic modelers is that a well-chosen numerical example may come close enough to more detailed analyses to help inform decisions that have to be made in hours during a crisis, not days or weeks. The insight that baseline vaccine capacity generates tens of trillions of dollars of surplus in a pandemic and that accelerating and expanding capacity is worth trillions more emerges as clearly from the numerical example as from the more detailed analysis. That said, toy examples are not perfect substitutes for methodical analysis. Indeed, the two sets of results in Table 1 are not identical: for instance, baseline capacity at risk avoids $17.0 trillion of harm in the numerical example but an estimated $18.8 trillion in Castillo et al. (2021). The $1.8 trillion difference is small in percentage terms but not surplus one would want to ignore in absolute terms.

3. Commercial Versus Social Incentives

The estimated value of baseline capacity from Castillo et al. (2021) is enormous ($6,200 per annual course), as is the estimated gain from investing in this capacity at risk ($1,000 per annual course) and the gain from doubling capacity ($870 per annual course added). As we saw, the enormous size of these estimates is not sensitive to detailed modeling assumptions: estimates from the numerical example are likewise enormous. Whichever estimates we focus on, they are several orders of magnitude larger than the price that manufacturers received during the pandemic for vaccines, between $6 and $40 per course (Castillo et al. 2021), suggesting an enormous gulf between social and commercial incentives for capacity investment.

To be sure, price is not a perfect measure of commercial incentives for investing in capacity. If anything, price is likely to overestimate commercial incentives. Budish and Snyder (2021) make the point in a numerical example in which a firm that is contracted to supply enough courses to vaccinate a billion people at $40 per course. The firm earns the same $40 billion whether it supplies the courses in one month or one year. The former rate requires 12 times the capacity investment, an expense which the firm would prefer to save if it does not increase its revenues. Whatever the level of the fixed price—whether in the $6 to $40 observed in the market for Covid-19 vaccines, even a figure approaching the thousands of dollars of value for accelerating capacity via at-risk investment—the fixity of the price eliminates much of the incentive for speed. With multiple competing vaccines, one firm may want to expand capacity to divert sales toward itself that would have later gone to
competitors, but such “business-stealing” incentives for capacity (Mankiw and Whinston 1986) may not be very strong in a pandemic with limited industrywide supplies.

Fixed prices also do not provide much incentive to accelerate capacity via at-risk investment. The probability that a vaccine succeeds through all the stages from phase-1 through phase-3 clinical trials on its way to winning approval is notoriously low. For example, Ahuja et al. (2021) took the probability of success for even the most promising candidate in their sample of Covid-19 vaccines under development to be less than 29%; other candidates in their sample had less than a 10% probability of success. Rather than sinking an investment that is likely to fail with limited ability to recoup if it does fail since the investment may not be very fungible, firms prefer to verify vaccines’ safety and efficacy in clinical trials first before investing in capacity at scale. If the firm receives the same price for the vaccine whether sold today or in three months, the firm has little incentive to invest before it more certain that the vaccine will succeed. By contrast, the social value of at-risk investment reported in the previous section was as high as $1,200 per annual course.

How can policy bridge the gap between commercial and social incentives for expanding and accelerating vaccine capacity? One possibility is to have buyers specify deadlines for delivery. Unless backed by penalties for late delivery, such deadlines may slip, a possibility highlighted by the suit filed by the European Union (EU) against AstraZeneca, which only delivered a third of the contracted 300-million doses by the specified date, citing production delays (Guarascio and Smout 2021). The EU sought damages of 10 euros per dose per day of delay; the judge instead relaxed AstraZeneca’s delivery deadlines and capped the penalty for delay at 10 euro per dose (not per day), far below the social cost of stalling the EU’s vaccination campaign, which depended heavily on AstraZeneca rather than alternative Covid-19 vaccines. Even if the judge had assessed damages on the order of the EU’s social loss in the billions or trillions of dollars, AstraZeneca would go bankrupt before paying the full amount, dulling incentives to avoid delay.\(^4\) Contracts can specify bonuses for early delivery, but the bonuses would have to be huge to have firms internalize the full social value of acceleration, and increase the danger that firms sacrifice quality to achieve early delivery.

Other approaches to bridge the gap between commercial and social incentives may be more promising. Rather than contracting on doses, buyers can contract directly on dedicated capacity in advance of investment. Contracting on doses may just result in the firm’s fulfilling orders in

\(^4\)In the law and economics literature, this is called the problem of the judgment-proof firm, first formally analyzed in Shavell (1986).
a slow trickle with the capacity it has. The buyer may find itself at the end of a long queue of customers, just as the EU found itself for AstraZeneca supplies behind the United Kingdom (UK), AstraZeneca’s home country. Instead of countries playing a Prisoners Dilemma for scarce supplies, involving negative externalities, contracts over capacity can generate positive externalities: once the dedicated capacity serves the contracting country, it can then be used to serve the needs of other countries. Rather than fighting over a position in the vaccine queue as entailed by contracts over doses, contracts over capacity can speed up the rate at which the queue moves.

It must be recognized that the situation may not be so sanguine if inputs into capacity expansion are severely limited. In that case, contracts over capacity could just entail shifting the Prisoners Dilemma back a stage, from fighting over scarce doses to fighting over scarce capacity. If inputs are truly this scarce and the supply chain this delicate, arguments for expanding input capacity, stockpiling storable inputs, and increasing the resilience of the supply chain in advance of future pandemics hold with even more force.

Another approach is to directly defray the firm’s capacity expenditures with “push” funding. As discussed in more detail in Section 6, “push” funding pays for a firm’s inputs whether or not a successful product is developed; this is distinct from “pull” funding that provides payments only on successful delivery. Push funding is particularly useful to induce at-risk investment. The firm is more willing to invest before it learns whether the vaccine is successful if the buyer shoulders most or all of the risk. Push funding is not without problems: it requires some degree of accounting control, provides little incentive for firms to economize on investment expenditures and improve their probability of success (moral hazard), and does little to discourage firms with no realistic chance of success from joining the market (adverse selection). A fuller discussion of the relative merits of push versus pull funding are left to Section 6. For now, we consider a funding mechanism involving a mix of push and pull, perhaps covering some percentage (say 85%) of the firm’s investment cost with push funding, leaving firms with some “skin in the game” to avoid some of the worst of the moral-hazard and adverse-selection problems. Push funding can be complemented with pull funding in the form of a purchase guarantee for successful doses lucrative enough to induce firms to contribute their portion of investment expenditures.
4. Optimal Vaccine Portfolio

Vaccine capacity can be worth trillions of dollars as we have seen, justifying billions of dollars of public spending to expand and accelerate it. Does this mean there is no end to the spending that can be justified on every, even remotely plausible, vaccine candidate? There is enormous value of at-risk investment in accelerating vaccine availability, but by definition this has to be undertaken when there is considerable uncertainty about which vaccines will succeed, especially early in the pandemic when there is the greatest opportunity to accelerate subsequent capacity.

This section explores approaches to determining the optimal portfolio of at-risk investment. We will show that the level of spending entailed, while bounded, is large enough to cause possible sticker shock. We will analyze which vaccine candidates are optimally included in a portfolio receiving at-risk investment depending on their stand-alone probability of success and correlation with other candidates using more or less similar technologies.

We focus on the methods formalized in Ahuja et al. (2021) but first used in the analysis informing the Athey et al. (2020) op-ed. Athey et al. (2020) was written prior to the launch of Operation Warp Speed by the United States (US) government, which ended up spending $18 billion on six Covid-19 vaccine candidates (Baker, Koons, and Kocieniewksi 2020). As “ambitious” and “unprecedented” as Operation Warp Speed was heralded to be (Slaoui and Hepburn 2020), the spending levels were an order of magnitude less than the recommendation in Athey et al. (2020) that the US spend $70 billion to develop and procure 15–20 vaccine candidates.

Athey et al. (2020) set up and analyzed the selection of the optimal portfolio of vaccines to receive funding from candidates then under development, in essence the same problem confronting officials in charge of programs like Operation Warp Speed at that time, a few months into the pandemic, when the first vaccine candidates were showing promise in early clinical trials. The authors considered a program with the following design. Each selected vaccine candidate is directed to invest at risk (in parallel with clinical trials) in enough capacity to fully vaccinate the entire population in the US deemed eligible for vaccination, requiring, say, 250 million courses, over a period of, say, three months. The program covers 85% of the at-risk capacity investment with push funding, leav-

---

5 The required capacity is larger than at first seems since 250 million courses of a vaccine requiring two doses for full vaccination amounts to 500 million doses, and capacity for 500 million doses in three months translates to annual capacity for 2 billion doses.
ing the remainder for the firm as “skin in the game.” Successful vaccines are promised a price per dose that builds in a markup over production cost calibrated to induce the optimal number of candidates to participate. The price, which serves as a sort of bonus, has to be high enough to allow the marginal participant to cover its 15% share of at-risk investment in expectation, taking into account the marginal candidate’s potentially low probability of developing a successful vaccine. For example, if the marginal candidate’s probability of success is 20%, then the bonus price would need to be five times its per-unit contribution to at-risk investment. This bonus price is a pull-funding component of the program, paying participants—not for inputs—but for successful outputs, providing incentives to mitigate moral-hazard and adverse-selection problems.

The optimal number of participants from the policymaker’s perspective maximizes program benefits minus costs. Program benefits involve some uncertainty, depending on whether any at-risk investments succeed. Given that each firm invests in enough capacity to vaccinate the country, expected benefits equal the probability at least one vaccine succeeds times the gain from successful at-risk investment. Assuming that the country is eventually fully vaccinated with or without the program, just more rapidly with the program, the gain from successful at-risk investment is the reduction in harm associated with the acceleration in vaccination embodied in region C of Figure 1.

To estimate the probability of success in a portfolio of vaccines accounting for possible correlation in outcomes between more similar vaccines, the authors developed a model specifying that a vaccine candidate fails if it experiences a problem in any one of several levels. Vaccines might not have worked for Covid-19 if it turned out to be as difficult a disease as malaria. Even if not, certain platforms might not result in successful vaccines—a particular concern with mRNA and DNA platforms, never before used for a vaccine in production. Within a platform like viral vectors, perhaps adenovirus would prove an unsuitable vector and only vaccines based on the measles vector viable. The fact that failure at one level causes all candidates branching out from that level to fail generates a positive correlation in failure among candidates. Just as negative correlation in returns is desirable in a portfolio of assets, negative correlation in failure is likewise desirable in a portfolio of vaccines, in this case increasing the probability of at least one success.

The authors based their estimates of these failure risks on a variety of academic sources, sup-

---

6 The bonus price would need to be yet higher if it was required to reflect a positive cost of capital. Pushing in the opposite direction, toward a lower required bonus price, is the fact that realized capacity from at-risk investment reduces the amount of additional investment a firm needs to reach its ultimate target capacity.
plemented with discussions with medical and industry experts. Even with careful modeling and estimation, substantial uncertainty remains in estimating success probabilities. Different probability estimates will select different portfolios as optimal. The truly optimal portfolio should be based on the most accurate probability estimates possible given available information. Improving the probability estimates may contribute as much to the surplus from the program as improving the selection mechanism given those probabilities.

Table 2 shows a list of the top candidates for the vaccine portfolio from the larger list considered by Ahuja et al. (2021). The table reports several relevant probabilities, both the candidate’s stand-alone probability of success as estimated by the authors as well as the candidate’s incremental contribution to the probability of at least one success in a portfolio selected by the “greedy” optimization algorithm. The greedy algorithm sequentially selects candidates from the remaining set that contribute the most to the criterion, in this case the probability of at least one success in the portfolio. The candidates in the table are listed in order of selection by the greedy algorithm. This ranking does not exactly correspond to the order of their stand-alone probabilities of success. For example, the algorithm selects the mRNA candidate before the inactivated-virus candidate appearing fourth despite the higher probability that the mRNA platform fails because a candidate using the inactivated-virus platform was already part of the portfolio.

Figure 2 shows the probability of at least one success rising in the size of the optimal vaccine portfolio. The curves’ concavity results from the concavity of the probability function and from the addition of less promising candidates as the portfolio grows. The black curve representing the selection by the greedy algorithm can be compared to the gray curve representing the portfolio from a naïve algorithm that selects candidates in order of their stand-alone probability of success. The curves are fairly close together, suggesting that their overall shape is not due to the sophistication of

---

7 Athey et al. (2020) took the perspective of an official designing a program as of that writing, in April 2020. Ahuja et al. (2021), written later, took the perspective of an official designing a program in August 2020, by which point the list of vaccine candidates in development had grown.

8 The use of different technologies and inputs by different vaccine platforms may provide another reason for preferring a diversified portfolio of candidates. A diversified portfolio may have a more elastic supply curve, lowering program costs. Ahuja et al. (2021) assume perfectly elastic capacity supply at $10 per annual course as a baseline. They explore alternative supply elasticities in Figure A7 of the online appendix, leading to different capacity prices analyzed in Table A8. However, they do not consider platform-specific supply elasticities, thus abstracting from cost reductions resulting from diversification.
the algorithm.\textsuperscript{9,10}

Combining the estimates of the probability of success with estimates of the gain to successful at-risk investment provides the needed information to compute expected benefits. It remains to choose the optimal number of candidates in the portfolio, maximizing expected program benefits minus costs, requiring consideration of the cost side. Expected program costs equal the program’s push funding for its share, say 85\%, of at-risk capacity investment for portfolio vaccines plus the bonus price for courses purchased by the pull component of the mechanism in expectation. The bonus price is not given but depends on which vaccines are brought into the portfolio. The bonus price has to be sufficiently lucrative to allow the marginal entrant to cover the capital costs for its share of at-risk investment or else it will not enter. The marginal entrant with the lowest probability of success discounts the bonus payment the most in computing its expected revenue from the bonus.

The point can most easily seen by considering specific portfolios selected from Table 2. The optimal five-candidate portfolio (i.e., consisting of the vaccines in the first five rows) would need to induce the candidate in the fifth row, with a 18.4\% probability of success, to enter. To pick some round numbers, suppose capacity costs $10 per annual course, with 85\% covered by push funding from the program, leaving the firm to cover the remaining $1.5 billion capacity investment. Assuming prices drop to production cost outside of the bonus period, so that the only source of profit is the bonus price, and taking the bonus price to be paid on the first 250 million accelerated courses, one can compute that the bonus price has to be $33 for the firm to break even. For the optimal 15-candidate portfolio, the marginal candidate has only a 9.2\% probability of success. Since this is half the marginal candidate’s probability of success in the five-vaccine portfolio, the bonus price must be doubled, to $66, to induce the marginal candidate to enter in this larger portfolio. Figure 3 incorporates these calculations into the overall calculation of expected program costs: $53.2 billion for the five-candidate portfolio, $177.0 billion for the 15-candidate portfolio.\textsuperscript{11}

\footnotesize
\textsuperscript{9}While appearing small to the naked eye, the gap between the curves is as large as two percentage points for a four-candidate portfolio. This gap represents billions of dollars of potentially lost surplus from an inferior algorithm when scaled by the trillions of surplus gained if at-risk investment succeeds.

\textsuperscript{10}The greedy algorithm provides a practical solution to NP-hard dynamic-programming problems but does not guarantee an optimal solution since it only looks ahead one step fixing the portfolio constructed up to that point, rather than fully re-optimizing. Using a brute-force search over all combinations of vaccines, in new analysis for this paper, we verified that the greedy algorithm could not be improved upon for portfolios of 1–7 vaccines.

\textsuperscript{11}A variety of additional considerations would lead the bonus price to have to be adjusted in one direction or the other. For simplicity, we have abstracted from those considerations in the text but provide a more detailed discussion in the appendix.
While the extra $123.8 billion in expected program cost for the 15-candidate over the five-candidate portfolio may seem hard to justify, it is justified according to estimates of expected program benefits above. As shown in Table 2, the probability of at least one success in the 15-candidate portfolio selected by the greedy algorithm is 15% higher than the five-candidate portfolio. As long as the gain from a three-month acceleration of vaccinations is at least $825 billion, a 15% higher chance of that amount justifies the extra program cost. From Table 1, we see that the estimated gain is about double this threshold ($1.6 trillion in the detailed calculations, more in the numerical example).

For pedagogical purposes here, we compared just two portfolio sizes. Ahuja et al. (2021) performed a more systematic comparison of all portfolio sizes. They refined the application of the greedy algorithm, adding candidates’ capacity to the portfolio in small increments rather than chunks sufficient to vaccinate the US population. They also applied the analysis not just to the US but to every country in the world.

Table 3 reports their results. From perspective of a program being designed from scratch based on the vaccine candidates available as of that writing (August 2020), the authors calculated that the optimal portfolio for the US would have included 27 candidates and at-risk investment in capacity for 462 million courses monthly. Some of this investment ends up being in failed candidates; in expectation, effective capacity for 98 million courses monthly is realized. Expected costs of the at-risk investment program are substantial, at $169.8 per capita, but are eclipsed by expected benefits, at $923.4 per capita. Optimal portfolios for countries that are smaller and/or poorer than the US involve fewer candidates but still a substantial number. For example, the optimal portfolio for Germany, the UK, and Australia includes 21 candidates and for Chile includes 12. The optimal portfolio averages 18.3 candidates across high-income countries, 6.7 candidates across middle-income countries, and 1.3 candidates across low-income countries.

As a result, different candidates end up installing different at-risk capacities in the optimal portfolio, with more promising candidates installing more capacity. As appropriate, their application of the greedy algorithm maximizes expected benefits minus costs directly rather than maximizing expected benefits first and later calculating the expected cost of resulting portfolios, as done here to simplify the exposition.

That some candidates invest less capacity than needed to serve the population in Ahuja et al. (2021) implies that there is a gain from having multiple successes in their analysis, unlike the illustrations in Figure 3 in which having one success is all that is needed to generate the full benefit. Multiple successes would also be beneficial in a model in which successful candidates have different efficacies. Even if efficacy differences are ex ante unpredictable, with multiple successes the program could roll out the one having the best efficacy draw.
5. Comparing Optimal Capacity in the Model to Actual

Comparing optimal capacity in the model to actual policy is difficult given data limitations. It is difficult to compare optimal to actual at-risk investment because capacity investments for failed candidates are seldom compiled, so actual at-risk investment is difficult to observe. We could try to set aside investments in failed candidates and just compare realized capacity for successful candidates across optimal and actual policies. The difficulty there is that the entry in Table 3 for realized capacity is a counterfactual expectation, which may differ from actual because of errors in forecasting probabilities of vaccine success.

A bound on the difference between optimal and actual capacity can be obtained by comparing optimal to actual capacity for just the subset of vaccines that were ultimately approved (globally, AstraZeneca, Pfizer, Moderna, Sinovac, and Sputnik). This is the approach taken in Ahuja et al. (2021). They estimated that if the world had invested at the recommended levels, it would have had capacity for 7.12 billion annual courses of approved vaccines rather than the three billion estimated from actual production runs, adding $1.7 trillion to benefits. This is undoubtedly an underestimate of the gains from investing at-risk since it does not reflect capacity from vaccines that would have been successful had they received investment.

The relatively small size of the optimal portfolio for low-income countries in Table 3 (averaging 1.3 candidates) and low associated net benefits (about 30 cents per capita) can rationalize why most of them did not secure bilateral deals for at-risk capacity early in the pandemic. Yet the vaccination rate in low-income countries remained low deep into the pandemic (Agarwal and Reed 2022). While Ahuja et al. (2022) can rationalize why low-income countries did not invest at risk, it does not rationalize persistently low vaccine purchases. The tiny benefit of an at-risk investment program to the average low-income country estimated by Ahuja et al. (2021), 30 cents per capita, vastly underestimates the gain from eventually purchasing vaccines there. The benefit estimate for at-risk investment reflects a very small portfolio of vaccines, any of which may fail, and only for a three-month acceleration of supply. The gain from purchasing vaccines after they have proven to be successful relative to not purchasing at all, measured on a per-course rather than per-capita basis, is orders of magnitude higher. For example, in new runs of the model of Castillo et al. (2021) for this paper, we find that sufficient capacity to vaccinate 40% of the population in every low-income
country within a year (allocated to the highest-value consumers in each country) would have been worth $350 per annual course at the start of 2021.

As discussed in Agarwal and Reed (2022), it may be hard to ever determine whether supply or demand factors bear more responsibility for the low vaccination rates in low-income countries. Low-income countries seem to have suffered lower mortality losses throughout the pandemic, perhaps due to their younger populations. The lack of orders in Africa’s first vaccine factory (Chutel 2022) may suggest that low demand rather than low capacity is currently limiting vaccinations there. That demand for vaccines in low-income countries is waning as the pandemic wanes does not prove that vaccine demand was low there at the outset of the pandemic. Agarwal and Reed (2022) suggest that greater demand for vaccines would have materialized in low-income countries early in the pandemic had loan programs been available earlier. Supporting this contention, surveys conducted by Solís Arce et al. (2021) found higher willingness to take Covid-19 vaccines in LICs and LMICs than some HICs.

High-income countries did not face the same financing constraints as low-income countries. What other explanations can be offered for their falling short of recommended investment levels? One explanation is based on the mathematics of optimization, which implies that slightly under- or overshooting an optimum does not entail much loss. If the US cut the budget of the at-risk program recommended in Table 3 in half, it would have obtained less net surplus than in the table, but not much less (91% of the full-budget net surplus according to Ahuja et al. 2021, Table A3). Even the half budget more than doubles total vaccine spending under Operation Warp Speed and more than doubles the number of vaccine candidates funded, implying that even the relatively generous program substantially undershot the optimum.

It is possible that the recommendations were themselves too high, not that actual investment was too low. The optimal portfolio calculations assumed that one dollar of expenditures costs one dollar. To the extent that the expenditure must be paid with taxes involving distortions to the markets in which they are raised, the calculations may have placed insufficient weight on saving money. On the other hand, officials may have placed undo weight on factors having more of a personal political cost than a true social cost such as avoiding being seen as giving handouts to big pharmaceutical firms or foreign suppliers, as backing projects that fail, or as wasting money on products going unused in a suddenly waning pandemic.
The recommended portfolios may have been too large if they were based on forecasts of probabilities of candidate success that were too pessimistic or correlations in candidates’ success that were too low. Those forecasting errors would have biased the recommendations toward large portfolios as the only way to obtain a reasonable probability of at least one success. Perhaps officials had better “inside” information that certain candidates were very likely to succeed and others not, allowing them to concentrate investment on a few candidates. Consistent with this possibility, of the six candidates funded by US Operation Warp Speed, four eventually obtained emergency use authorization from the US government and a fifth obtained a similar designation from the EU, a remarkable 83% success rate, almost three times higher than the highest estimated probabilities in Table 2. As mentioned above, the analysis in Ahuja et al. (2021) suggests it would have been worthwhile to have invested in more at-risk capacity for those candidates even without such “inside” information. Having “inside” information on which candidates to concentrate would only strengthen the case for expanding at-risk investment in those candidates.

Political economy may provide another explanation. The international crisis demanded action, but it may have been unclear who would take the lead and how investments would be paid for. Perhaps countries were waiting to see which multilateral organization would step up to design and fund an effective vaccine program, delaying countries’ response.

Another explanation is that the main points in this paper were insufficiently appreciated by policymakers. Ironically, those with the deepest expertise in optimal vaccine policy in ordinary times may have had the hardest time switching to the crisis-mode mindset befitting a pandemic. Policymakers may have understood that large investments were urgently needed but not appreciated just how large. They may have regarded capacity as more fixed, less responsive to more spending, than it truly was. They may not have appreciated that incentives for capacity expansion are eroded by a fixed price per dose. If we can clarify the principles now, they may become part of conventional wisdom guiding the response to the next pandemic.

6. Further Details on Program Design

This section elaborates on a series of design features that can improve the efficiency of programs to accelerate vaccine supply.
6.1. Push Versus Pull

The relative merits of push versus pull funding, previously hinted at, requires more discussion. The terms push and pull mean different things to different authors. In the analysis in Section 4, the terms were used in their relatively pure forms. Push funding meant the act of defraying a supplier’s documented investment costs (say cost of clinical trials or capacity installation) as they are expended before learning of the project’s success. Pull funding meant a contract to purchase courses of a supplier’s approved vaccine (or a broader commitment to buy approved vaccines from any qualified manufacturer).

Push funding can help economize on program expenditures. With another party picking up the tab for its investment costs, there is no reason for a firm not to invest. The program only needs to cover the total investment cost of participating firms (perhaps a margin above that to cover risk-adjusted capital costs). There is no need to provide a premium to more or less productive firms. The pull funding considered in Section 4 came in the form of a fixed-price rather than cost-plus contract. The price needed to induce the marginal candidate to enter is higher than needed to induce more productive inframarginal candidates to enter. Absent an ability to price discriminate, the high marginal price has to be paid to all candidates, resulting in higher program expenditures under pull funding than push funding.

To take a concrete example, consider the five-candidate portfolio in Figure 3. The funding mechanism shown, which covers 85% of firms’ capacity investment with push funding and the rest with pull funding entails expected total expected expenses of $53.2 billion. The $33 price needed to induce entry of the marginal firm provides a rent to the four inframarginal firms since, because of their more likely success, they would be willing to enter even for a bonus price around $5 to $10 less per course. This “overpayment” to inframarginal firms raises program expenses. A pure push mechanism covering 100% of the firms’ capacity investment would only cost $50 billion, a $3.2 billion savings over the hybrid mechanism in the example.

If the mechanism could discriminate by offering different pull prices to different firms, pull funding could be as economical as push. However, legal or informational constraints may prevent discriminatory price offers. The firm’s probability of success may be its private information, preventing the program from being able to condition price on probability of success.

Push funding has drawbacks, explaining why Ahuja et al. (2021) advocated a hybrid mecha-
nism with some pull to complement the push funding. Reimbursing all the firm’s costs regardless of success exacerbates moral-hazard problems, giving the firm little incentive to take measures to improve the probability of success, halt unpromising projects, economizing on costs. Push funding also exacerbates adverse-selection problems, doing little to keep “hobbyist” entrants with unrealistic prospects out of the portfolio and little to encourage promising entrants unknown to program, perhaps in lines of business outside vaccines, to identify themselves. For those reasons, Ahuja et al. (2021) recommend a funding mechanism that is mostly push but leaves firms with enough “skin in the game” to mitigate moral hazard and adverse selection.

6.2. Bonus for Speed

A further advantage of pull funding is that it can be structured to incentivize speed. Indeed, the example in Figure 3 builds in such a structure: the price in the pull-funding component is provided as a bonus payment for courses produced on an accelerated schedule. Though the bonus pales in comparison to the social value of acceleration, it reflects a substantial return on investment compared to the firms’ small (15%) contribution to capacity costs and may be sufficient to induce the firm to take most reasonable steps toward accelerating production.

6.3. Private Information

The solution for the optimal vaccine portfolio in Section 4 technically assumed that the designer had perfect information about the firms (although could not exploit the information to discriminate on the bonus price offered in the pull-funding component). Snyder, Hoyt, and Gouglas (2022) solve for the optimal funding mechanism when firms’ productivity is private information.14

A companion paper by Snyder et al. (2020) applied this general mechanism to the procurement of Covid-19 vaccines assuming firms’ private information regards their clinical trial and capacity costs. Using proprietary data from a survey of potential entrants in the market for Covid-19 vaccines, they estimated the investment-cost distribution from which potential entrants draw. That, together with estimates of the benefits of vaccine capacity, were used to calibrate the optimal mechanism. Despite

---

14The mechanism-design problem involves some non-standard elements including that multiple units need to be procured from multiple suppliers with failure risk. The optimal mechanism turns out to be a \( w^* + 1 \) price Vickrey auction with reserve, where the number of winning firms \( w^* \) maximizes the virtual surplus function.
differences in perspective (perfect versus private information), models, and model inputs, calibrations of program costs and benefits in Snyder et al. (2020) are similar to comparable calibrations from and Ahuja et al. (2021), providing confidence in the robustness of their conclusions.

6.4. International Coordination

At the outset of the Covid-19 pandemic, initiatives to coordinate vaccine procurement in a single international body coalesced in the formation of COVAX, lead by Gavi—the Vaccine Alliance, the Coalition for Emergency Preparedness Innovations (CEPI), and the World Health Organization (WHO). Proposed benefits to COVAX participation included access to a larger and thus more diversified portfolio of vaccine candidates, lower prices due to enhanced bargaining power, and coordinated donations to LICs, which might not be able to afford vaccines on their own (Berkley 2020).

While high-income countries signed up with COVAX, most ultimately procured the bulk of their vaccines outside of the program, through direct deals with manufacturers. We may never learn how much more at-risk capacity would have resulted from deeper international cooperation. On the one hand, most countries fell short of the recommendations in Ahuja et al. (2021) for their stand-alone optima—to say nothing of the joint optimum taking into account epidemiological externalities across countries, presumably involving higher investments. On the other hand, it is theoretically possible for intense competition for capacity in bilateral deals to generate more capacity than a coordinated international program (Snyder et al. 2020), certainly compared to an underfunded version of the program.

Some proposed benefits of international centralization emerged despite the prevalence of decentralized, bilateral deals. The US, EU, and UK, among other large, high-income countries, managed to diversify their vaccine portfolios via bilateral contracts. Prices did not spike during the pandemic as some had predicted despite the intense shortage and competition for doses. However, the decentralized approach has suffered from some shortcomings, the most glaring according to some contemporaneous commentaries (e.g., Mueller and Robbins 2021) being the slow vaccine rollout to LICs and MICs. With better funding, COVAX could have accelerated supplies to these countries, not only by bidding more aggressively for a place in the queue for existing vaccine capacity but also by bidding to expand that capacity. Castillo et al. (2021) estimated that investing in an additional billion annual courses of capacity even at the late date of that writing could have accelerated com-
pletion of a vaccination campaign in LICs by as much as 10 months. In more recent experience, despite an increase in vaccine supplies to LICs, coverage there seems to be asymptoting well short of the 70% threshold considered in Castillo et al. (2021), suggesting that more capacity may not have substantially accelerated vaccinations there. On the other hand, it could be argued that more people in LICs would have taken vaccines if they were available earlier, when Covid-19 was perhaps seen as a bigger threat. This argument is supported by survey evidence from Solís Arce et al. (2021) of a higher willingness to take Covid-19 vaccines in LICs even compared to the US among other HICs.

A centralized procurement mechanism cannot conscript sovereign countries to participate; country participation must be incentive-compatible. Participation is a particular question for high-income countries, which have the best options to pursue bilateral contracts outside of the centralized mechanism, as many did outside of COVAX. In their calibrated model, Ahuja et al. (2021) found that HICs would opt out of a centralized mechanism allocating vaccine according to population but exacting contributions proportional to gross domestic product per capita. A centralized mechanism can only sustain so much cross subsidization before collapsing. The scope for cross subsidization is greater the stronger are cross-country epidemiological externalities and the more power the centralized mechanism has to bargain for lower procurement prices.

Short of fully cooperating in a global procurement mechanism, countries could help advance global welfare by at least refraining from taking the deleterious step of imposing export controls, allowing firms to ship vaccine to foreign countries when domestic needs are fulfilled and sufficient stockpiles have accumulated. Export controls can be particularly damaging if imposed on essential inputs into vaccine production. Bown (2022) documents the complexity of vaccine supply chains, with inputs from several countries often needed to produce finished vaccines in another. Restrictions that gum up these supply chains can have a cascading effect, reducing finished vaccine supplies everywhere. Global welfare would likely be raised by removing such restrictions, a straightforward step which might not require undue international coordination. Global welfare would likely be further increased by further cooperative efforts to incentivize expanded capacity of inputs along the supply chain that might form bottlenecks, following many of the same principles as described here for finished vaccine production.15

---

15Bown, Snyder, and Staiger (2022) provide a theoretical analysis of selected international trade and supply-chain considerations associated with the Covid-19 pandemic, cataloged comprehensively in Agarwal and Gopinath (2022).
Whether the global allocation of vaccine is centralized or not, it could be improved if complemented by a cross-country vaccine exchange (Budish, Kominers, and Prendergast 2022). For example, a country whose COVAX allocation includes several vaccines may prefer to simplify logistics by consolidating on one or may prefer to trade for a vaccine that does not require cold storage facilities that it lacks. A country holding onto a stockpile in case boosters are needed later might be willing to make an intertemporal trade with a country that has secured delivery in future months but needs vaccines now. A high-income country with excess supplies could use the exchange to gift those supplies to needy countries in a systematic way.

6.5. Contracting on Capacity Versus Doses

The danger in contracting on a number of vaccine courses to be delivered is that the supplier can fulfill the contract by installing a small amount of capacity and producing the desired quantity slowly over time, earning the same contractual revenue (if the deal involves fixed prices) while economizing on capacity expenditures. Any incentives that the buyer provides for timely delivery may result in moving it up the queue by pushing other customers back.

Directly contracting on capacity can mitigate these problems. The buyer guarantees itself first dibs on the output of that capacity. Furthermore, the expanded capacity can then be switched to other buyers after serving the first’s needs. Instead of lengthening the vaccine queue—a negative externality on other buyers—a contract on capacity can increase the rate at which the queue is served—a positive externality. If inputs needed to expand capacity are perfectly inelastically supplied, then conflicts and negative externalities are unavoidable: moving from contracts on courses to contracts on capacity simply shifts competition over courses to competition over capacity. To the extent capacity installation is elastic, however, contracts on capacity are more likely to generate positive externalities.

7. Dose Stretching

Given the tremendous value of scarce vaccine capacity during a pandemic, various policies to stretch existing capacity are worth exploring and exploiting. One such policy, labeled “first doses first,” delays the second dose of a two-dose sequence so that more people can obtain at least one shot
more quickly. There is precedent for the use this form of dose stretching. The UK resorted to a first-doses-first policy in December 2020, while still facing supply constraints in the early stages of its vaccination campaign against Covid-19, extending the interval between vaccine doses from four weeks to twelve weeks. A number of other countries followed suit (Cortez 2021).

Capacity can also be stretched by “fractional dosing,” reducing the amount of active ingredient in each dose. To the extent that the active ingredient is the rate-limiting factor, not glass vials or other inputs, halving the amount of active ingredient in a dose can double output from a given capacity. If the fractional dose entails little loss of efficacy, the policy can be an obvious “win” for public health by relieving supply constraints. Even if efficacy is reduced substantially, fractional doses may still be a “win” for public health since the alternative is to leave more people completely unprotected for longer. Fractional dosing has been employed successfully in historical examples, including the 2016—18 outbreak of yellow fever in Africa, leading the WHO to recommend 1/5 doses to stretch available supplies (World Health Organization 2017).

In ordinary times, pharmaceutical firms could conduct extensive tests to determine the optimal amount of active ingredient in a dose, number of doses, and time between them, seeking potency while avoiding undue side effects and expense. In more urgent circumstances of a pandemic, having less time to perfect dosage protocols, firms may be inclined to satisfice, possibly erring on the side of potency if it is uncertain that vaccines against a novel disease will even work. However, in a context in which there are shortages of vaccines, erring in the opposite direction, toward lower potency, can improve public health by vaccinating more people quickly even at the sacrifice of efficacy for individuals.

For Covid-19 vaccines, evidence from clinical trials suggests that lower doses of certain vaccines produce strong immune responses (as measured by neutralizing antibodies), especially against the original strain of the virus. Combined with early research on correlates of immunity, this evidence suggests that 1/2 or even 1/4 doses of some vaccines could be highly effective—possibly even more effective than full doses in some cases (Więcek et al. 2022).

In epidemiological modeling conducted by Więcek et al. (2022), fractional dosing was found to avert more death than full dosing across a range of different scenarios, those assuming that fractional dosing results in a substantial efficacy drop. For example, in baseline simulations (with an effective reproductive number of 2, an initial infection rate of 0.1%, and 70% efficacy for a full dose), moving
from full doses to 1/2 doses that are 80% as effective averted a third more deaths because of the offsetting acceleration in population coverage. Full doses averted the same number of deaths as 1/2 doses only when the efficacy of 1/2 doses dropped to 40% that of full doses.

Suggestive evidence of this sort, no matter how promising, may not be enough to convince authorities to upend the status quo and implement dose-stretching measures without direct testing of those measures in randomized controlled trials (e.g., World Health Organization 2021). That is no argument for inaction. According to the suggestive evidence, chances are good that dose stretching would be socially valuable; however, clinical trials of dose stretching would be worth undertaking even if those chances were remote. Clinical trials typically cost in range of millions or tens of millions of dollars (Gouglas et al. 2018), while the social value of capacity can be in the trillions.

Given the enormous potential benefits of dose stretching and the relatively small costs of clinical trials, experimenting with dose stretching for Covid-19 vaccines is an investment with very high expected social returns. However, firms’ private returns from these experiments is much lower than the social return. The ultimate goal of such experiments are measures to accelerate vaccination, and as we have repeatedly argued, firms’ private returns from accelerating vaccination are lower than the social returns. Experimentation has additional pitfalls for firms, who may be reluctant to expose their product to the risk clinical failure once a success has been registered, or to any negative efficacy news, even for dosing regimens that are not used, which may confuse the public.

This suggests a role for public investment in experimentation on dose stretching. For future pandemics, numerous dose sizes and dosing schedules could be tested from outset of vaccine development. Więcek (2022) provides more discussion of the relative merits of specific testing approaches and their optimal designs. Public institutions, which often subsidize vaccine research anyway, can provide further subsidies to expand clinical trials to optimize dosing regimens. In addition to, or instead of, randomized controlled trials, authorities could roll out dose stretching on a limited, experimental basis. If the observational evidence is positive, the dose-stretching program can be expanded. If negative, the program can be curtailed. Those receiving the stretched dose need not suffer permanent damage as they can be revaccinated with a full dose. This sort of experimentation has considerable option value: large upside potential with little downside risk.

Agarwal and Gopinath (2022) make the general case that research and development is a public good in a pandemic, calling for extensive public funding of clinical trials among other forms of research and development.
8. Conclusion

The Covid-19 pandemic wreaked nearly incalculable harm in every corner of the world map throughout all aspects of human life, causing sickness, death, and output losses; impairing people’s educations, career prospects, and daily social interactions. The world responded with unprecedented speed, mobilizing financial and institutional resources to address the challenges. An arsenal of new vaccines—some involving technologies (mRNA) never used for approved vaccines in human history—were successfully developed and administered to billions of individuals. As miraculous and heroic as the response was, there was scope for improvement. Calls for policies to accelerate and expand vaccine capacity that would appear outlandish during ordinary times made obvious economic sense in a pandemic.

The shortage of vaccine supply observed in many countries could have been mitigated by expanding at-risk investment early in the pandemic in a portfolio of vaccines, each with enough capacity so that much of the world could be rapidly vaccinated even if just one or a few end up successful. Firms lack the private incentives to invest at that speed and scale, so public policy must fill in to incentivize it. Advanced purchase mechanisms combining push and pull funding allow the risk of investing in projects like vaccines with low probability of success to be transferred from firms to funders at the same time building in incentives for speed and quality. The optimal level of public funding might cause sticker shock, but is eclipsed by potential benefits—spending billions to save trillions. To be sure, there may be physical limits to how much capacity can be expanded in a short time frame; but virtually any way to expand capacity with more resources could be exploited.

Policies to stretch scarce capacity also merit experimentation. The UK set an important precedent by quickly evaluating and approving the delay of the second vaccine dose to accelerate availability of first doses to more people. Lives were saved not just in the UK but in the countries that followed the UK’s example. Other promising dose-stretching strategies such as fractional doses were left mostly unexplored. Accelerating vaccination using such strategies could have averted substantial disease burden particularly in LICs and LMICs, which continue to face tighter supply constraints.

This article focused on the causal chain leading from (a) expanded vaccine capacity to (b) accelerated vaccinations to (c) mitigating pandemic harm. The logic is more general, applying beyond
expanding capacity for finished vaccine production to capacity for other inputs that would create a more resilient supply chain. The logic applies to mitigation measures beyond vaccines including personal protection equipment, drug treatments, and viral/antibody testing. The logic applies beyond investments accelerating the supply of physical products to investments accelerating the availability of information—including improved research and experimentation infrastructure, facilitating larger initial clinical trials of multiple dosage regimens, continuous trials of vaccines targeting evolving variants, and better monitoring of infection and mortality rates. The article focused not only on having more capacity but having it earlier by investing at risk, before clinical trials are completed. The logic of at-risk investment applies to any product or process that like vaccines involves a long “time to build” and considerable uncertainty regarding success.

The lessons learned about the value of vaccine capacity in the Covid-19 pandemic can help prepare for future pandemics. Since there are limits to how much capacity can be expanded during a short-run crisis period, governments and international organizations should consider investing in “peacetime,” in advance of the next pandemic. The goal should be to ensure that a global vaccination campaign can be completed within months—not years—the next time a pandemic strikes. To this end, policymakers could consider creating and maintaining sufficient manufacturing capacity to vaccinate a target proportion (say up to the threshold for herd immunity) in a target period (say six months) not merely with one vaccine but with redundant capacity for several candidates to allow at-risk investment and guard against the risk of failure for one or another. We do not know what the next pandemic pathogen will be, but repurposing existing capacity (especially if designed to a standard that facilitates repurposing) is much faster than building facilities from scratch. In addition, general-purpose inputs to vaccine production should be stockpiled ahead of time. As witnessed during Covid-19, the attempt to vaccinate billions of people all at once can quickly exhaust available inputs, which can only be slowly replenished with the pandemic disrupting supply chains. These measures will not come cheap, but if the next pandemic has anywhere near the health and economic costs of Covid-19, investments in building and maintaining capacity and stockpiling inputs will have high returns.

The extent to which countries played a zero-sum game with their responses to the Covid-19 pandemic can be debated. Contemporaneous reports decried the hoarding of medical supplies and restrictions of trade (Goodman, et al. 2002). At least 69 countries imposed some sort of restriction
on restricted medical exports early in the pandemic; the US, for example, sought to ban 3M from exporting masks even to its neighbors to the north and south. The policies covered in this paper could have helped avert a zero-sum game. Dose stretching and international cooperation are obviously positive-sum strategies. Bilateral contracts to accelerate and expand capacity are less obviously positive-sum, but they could be if done correctly: contracting on capacity rather than doses can increase the rate at which the country itself is served, at which point the expanded capacity can be turned to supply the rest of the world. Whatever one thinks the nature of the game is during a pandemic, outside of a pandemic, in peacetime, the game certainly need not be zero-sum. Because the supply of capacity is elastic in the long run over which peacetime investments are made, contracting to install additional capacity need not deprive other countries looking to make similar investments. Having adequate capacity in place before the start of a pandemic ameliorates the scarcity that drives hoarding, trade restrictions, and other zero-sum behavior during a pandemic.

A political constraint on peacetime investment is that policy interest naturally wanes as the pandemic wanes and other priorities compete for attention. It will be important to document high social returns to peacetime investment in pandemic preparedness relative to other priorities with solid economic analysis. The case will be more persuasive if it can be shown that returns are high even based on conservative estimates of the probability of future pandemics, as work underway indicates. The lessons may apply sooner rather than later: the “future pandemic” may not be a novel disease coming many years from now but another Covid-19 wave sparked by a new variant that escapes current natural and vaccine immunity. Our modeling suggests potentially high social returns from ensuring adequate provisions are made for this forecastable possibility. As with all the investments discussed in this paper, the loss function involved is asymmetric. The “waste” of investments made for a Covid-19 wave that does not materialize are eclipsed by the harm suffered if the wave crashes over a world that is inadequately prepared.

---

17 Building on estimates of the power-law distribution of historical pandemic intensity from Mariani et al. (2021), Glennerster, Snyder, and Tan (2022) estimate that expected global losses from pandemics exceed $300 billion annually. Investing $60 billion up-front in expanding capacity for supply-chain inputs and vaccine capacity and an expected annual flow of $2.2 billion in production costs thereafter could generate expected global net benefits of $28 billion annually.
Appendix: Complications to Bonus-Price Calculation

The bonus-price calculation illustrated in Figure 3 abstracts from a variety of complications, which this appendix returns to discuss. One complication is that, rather than just breaking even, to justify investment, firms need to earn a positive expected rate of return on invested funds, a rate which increases in the riskiness of the project. Taking this rate of return to be 20%—the mean for late-stage pharmaceutical projects from the survey by Villiger and Nielsen (2011)—the bonus prices in Figure 3 would have to be 20% higher to induce the marginal firm to invest, increasing the cost of the five-candidate portfolio by $2.1 billion and the 15-candidate by $9.8 billion.

A second complication, working in the opposite direction to make the at-risk portfolio look more attractive, is that any capacity realized from at-risk investment reduces the subsequent investment needed to reach a firm’s long-run target capacity. These indirect savings are not accounted for in the calculations presented in the text, either in calculations of the health benefit of at-risk investment (as in region C of Figure 1) or of at-risk portfolio costs (as in Figure 3), so requires separate accounting. In the illustrative example of Figure 3, expected indirect savings can be calculated as the product of the probability of at least one success in the portfolio and the avoided expense of the installation of one billion annual courses of capacity by a successful firm ex post. We calculate that this savings is $70.8\% \times $10\text{ billion} = $7.1\text{ billion}$ for the five-candidate portfolio and $85.8\% \times $10\text{ billion} = $8.6\text{ billion}$ for the 15-candidate portfolio. These indirect savings can be subtracted from direct portfolio costs in Figure 3 to determine net portfolio cost.

The two complications offset each other, sometimes almost exactly as in the 15-candidate portfolio. For simplicity, the illustrative example in Figure 3 abstracts from both complications, as do the calculations in Ahuja et al. (2021).

A third factor, also making the at-risk portfolio look more attractive, is that at-risk investment in a failed candidate may be repurposed to expand capacity for successful Covid-19 vaccines or other uses, reducing the penalty for failure. The illustrative example and Ahuja et al. (2021) both maintain the extreme assumption that capacity for a failed candidate is not fungible.
References


### Table 1: Global value of alternative vaccine-capacity policies

<table>
<thead>
<tr>
<th>Policy</th>
<th>Vaccine capacity (billion courses annually)</th>
<th>Global benefit (trillion $)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Numerical example</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Baseline capacity not at risk ((A+B))</td>
<td>3.2</td>
<td>14.0</td>
</tr>
<tr>
<td>• Baseline capacity at risk ((A+B+C))</td>
<td>3.2</td>
<td>17.0</td>
</tr>
<tr>
<td>• Double baseline capacity at risk ((A+B+C+D))</td>
<td>6.4</td>
<td>20.5</td>
</tr>
<tr>
<td><strong>Castillo et al. (2021) estimates</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Baseline capacity not at risk</td>
<td>3.0</td>
<td>15.8</td>
</tr>
<tr>
<td>• Baseline capacity at risk</td>
<td>3.0</td>
<td>18.8</td>
</tr>
<tr>
<td>• Double baseline capacity at risk</td>
<td>6.0</td>
<td>21.4</td>
</tr>
</tbody>
</table>

**Notes:** Letters in rows under the numerical example refer to regions in Figure 1. Baseline capacity in the numerical example computed as follows: 7.9 billion global population $\times$ 70% coverage $\div$ 21 months to reach target coverage $\times$ 12 months per year $= 3.2$ billion courses of annual capacity.

**Source:** Results for numerical example in the top set of rows derived from authors’ calculations based on Figure 1. In the bottom set of rows, result for baseline capacity at risk taken from Table S2 from the online supplement to Castillo et al. (2021). Remaining results in the last set of rows derived from new runs of the Castillo et al. (2021) model for this paper.
Table 2: Candidates for optimal vaccine portfolio

<table>
<thead>
<tr>
<th>Platform</th>
<th>Subcategory</th>
<th>Clinical stage</th>
<th>Probability of at least one success in portfolio (%)</th>
<th>Probability of individual vaccine success (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inactivated</td>
<td>Standard</td>
<td>Phase 3</td>
<td>28.8</td>
<td>28.8</td>
</tr>
<tr>
<td>Viral vector</td>
<td>Adenovirus</td>
<td>Phase 3</td>
<td>48.4</td>
<td>19.6</td>
</tr>
<tr>
<td>mRNA</td>
<td>LNP-encapsulated</td>
<td>Phase 3</td>
<td>58.4</td>
<td>10.0</td>
</tr>
<tr>
<td>Inactivated</td>
<td>Standard</td>
<td>Phase 3</td>
<td>65.8</td>
<td>7.4</td>
</tr>
<tr>
<td>Protein subunit</td>
<td>Recombinant</td>
<td>Phase 2</td>
<td>70.8</td>
<td>5.0</td>
</tr>
<tr>
<td>Protein subunit</td>
<td>S protein</td>
<td>Phase 2</td>
<td>74.5</td>
<td>3.7</td>
</tr>
<tr>
<td>Protein subunit</td>
<td>Recombinant</td>
<td>Phase 2</td>
<td>77.0</td>
<td>2.5</td>
</tr>
<tr>
<td>mRNA</td>
<td>LNP-encapsulated</td>
<td>Phase 3</td>
<td>79.0</td>
<td>2.1</td>
</tr>
<tr>
<td>Inactivated</td>
<td>Standard</td>
<td>Phase 3</td>
<td>80.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Viral vector</td>
<td>Adenovirus</td>
<td>Phase 2</td>
<td>82.1</td>
<td>1.4</td>
</tr>
<tr>
<td>Virus-like particle</td>
<td>Standard</td>
<td>Phase 1</td>
<td>83.3</td>
<td>1.2</td>
</tr>
<tr>
<td>Viral vector</td>
<td>Adenovirus</td>
<td>Phase 2</td>
<td>84.1</td>
<td>0.8</td>
</tr>
<tr>
<td>Viral vector</td>
<td>Measles</td>
<td>Phase 1</td>
<td>84.7</td>
<td>0.7</td>
</tr>
<tr>
<td>Protein subunit</td>
<td>S protein</td>
<td>Phase 1</td>
<td>85.3</td>
<td>0.6</td>
</tr>
<tr>
<td>DNA</td>
<td>Electroporation</td>
<td>Phase 2</td>
<td>85.8</td>
<td>0.5</td>
</tr>
<tr>
<td>Protein subunit</td>
<td>S protein</td>
<td>Phase 1</td>
<td>86.2</td>
<td>0.4</td>
</tr>
<tr>
<td>Live attenuated</td>
<td>Standard</td>
<td>Preclinical</td>
<td>86.5</td>
<td>0.3</td>
</tr>
<tr>
<td>DNA</td>
<td>Other DNA</td>
<td>Phase 2</td>
<td>86.8</td>
<td>0.3</td>
</tr>
<tr>
<td>Live attenuated</td>
<td>Standard</td>
<td>Preclinical</td>
<td>87.1</td>
<td>0.3</td>
</tr>
<tr>
<td>Protein subunit</td>
<td>Recombinant</td>
<td>Phase 1</td>
<td>87.3</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Notes: Candidates ranked in order of selection by greedy algorithm. Only top 20 shown of complete list considered by Ahuja et al. (2021) of all 142 candidates under development by a cutoff date (August 2020) determining the ex ante period. Based on discussions with experts and published estimates, the authors used the following inputs for the o-ring model of vaccine development: 10% chance of a problem at the overall level of Covid-19 vaccines, 20% chance of a problem in each platform (increased to 40% for mRNA and 60% for DNA because of their untried status), and 20% chance of a problem in each subcategory. Within a subcategory, individual vaccines fail with probability that falls the farther it is along in clinical development: 86% if in from the preclinical stage, 77% in phase-1, 68% in phase-2, and 50% in phase-3 clinical trials. “Standard” subcategory indicates that it is the sole subcategory for the platform; remaining platforms have more than one subcategory in complete list.

Source: Table A1 from the online appendix to Ahuja et al. (2021), supplemented by authors’ calculations.
Table 3: Results for optimal vaccine portfolios by individual countries and categories

<table>
<thead>
<tr>
<th>Country or category</th>
<th>Vaccine candidates</th>
<th>At-risk capacity</th>
<th>Expected program benefits</th>
<th>Expected program costs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total distinct (count)</td>
<td>Mean across countries (count)</td>
<td>Total installed (mil. courses per month)</td>
<td>Total realized (mil. courses per month)</td>
</tr>
<tr>
<td>Categories of country</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• All</td>
<td>30</td>
<td>8.8</td>
<td>2290.0</td>
<td>538.9</td>
</tr>
<tr>
<td>• High income</td>
<td>30</td>
<td>18.3</td>
<td>1418.0</td>
<td>308.0</td>
</tr>
<tr>
<td>• Middle income</td>
<td>18</td>
<td>6.7</td>
<td>906.9</td>
<td>239.3</td>
</tr>
<tr>
<td>• Low income</td>
<td>18</td>
<td>1.3</td>
<td>2.3</td>
<td>0.6</td>
</tr>
<tr>
<td>Selected individual countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• United States</td>
<td>27</td>
<td>—</td>
<td>462.3</td>
<td>98.0</td>
</tr>
<tr>
<td>• Germany</td>
<td>21</td>
<td>—</td>
<td>113.3</td>
<td>24.2</td>
</tr>
<tr>
<td>• United Kingdom</td>
<td>21</td>
<td>—</td>
<td>85.3</td>
<td>18.4</td>
</tr>
<tr>
<td>• Australia</td>
<td>21</td>
<td>—</td>
<td>34.5</td>
<td>7.4</td>
</tr>
<tr>
<td>• Chile</td>
<td>12</td>
<td>—</td>
<td>10.9</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Notes: Program benefits reflect the gains from accelerating capacity availability by three months, not the difference between having the vaccine portfolio and not. Entries for categories of countries represent the aggregation of individual country programs, not the optimal program for that coalition (although it may approximate the coalition optimum). Mean vaccine candidates omitted for individual countries since entry would be same as for total distinct. Per-capita at-risk capacity for low-income countries is positive (0.002) but rounds to zero in displayed entry.

Source: Table A2 from the online appendix to Ahuja et al. (2021), supplemented by authors’ calculations.
Figure 1: Example illustrating benefits from accelerating and expanding vaccine capacity

Notes: Diagram of numerical example in which world suffers $1 trillion harm each month from pandemic, eliminated once vaccinations cover 70% of population. The functional forms of all curves are dictated by the numerical example except for the boundary of $B$, which can be any convex function intersecting $A$'s lower boundary at its two endpoints. We take the curve to be the unique parabola that in addition to those properties has a vertex on the right intersection. Source: Authors’ calculations based on numerical example in text.
**Figure 2:** Probability of success from vaccine portfolios

Notes: Greedy algorithm indicated by black curve adds vaccine candidates to portfolio in order of increment to probability of at least one success. Naïve method indicated by gray curve selects candidates in order of individual probability of success, breaking ties randomly (curve is mean from 1,000 repetitions).

Source: Authors’ calculations using data from Table A1 from the online appendix to Ahuja et al. (2021). The black curve reprises Figure A6 from that source.
Figure 3: Computing program costs for several portfolio sizes

Five-candidate portfolio

Push-funding component:

5 candidates in portfolio \times 1 \text{ billion annual courses of capacity} \times $10 \text{ cost per annual course} \times 85\% \text{ program share} = \$42.5 \text{ billion}

Pull-funding component:

1.3 expected successes \times 250 \text{ million accelerated courses} \times $33 \text{ bonus price per course} = \$10.7 \text{ billion}

$53.2 \text{ billion}$

Fifteen-candidate portfolio

Push-funding component:

15 candidates in portfolio \times 1 \text{ billion annual courses of capacity} \times $10 \text{ cost per annual course} \times 85\% \text{ program share} = \$127.5 \text{ billion}

Pull-funding component:

3.0 expected successes \times 250 \text{ million accelerated courses} \times $66 \text{ bonus price per course} = \$49.5 \text{ billion}

$177.0 \text{ billion}$

Notes: The expected successes computed by adding up the individual probabilities of success over candidates in the portfolio. For example, in the case of the five-candidate portfolio, $28.8\% + 28.8\% + 21.6\% + 28.8\% + 18.4\% = 1.3$. Bonus prices rounded up to whole dollars. Remaining components explained in the text.

Source: Authors’ calculations.