

NBER WORKING PAPER SERIES

DESIGNING ADVANCE MARKET COMMITMENTS FOR NEW VACCINES

Michael Kremer  
Jonathan D. Levin  
Christopher M. Snyder

Working Paper 28168  
<http://www.nber.org/papers/w28168>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
December 2020

The authors served as members of the AMC Economics Expert Group that helped design the pilot pneumococcal AMC. They served as advisors to GAVI, Inter American Development Bank, and World Bank on funding mechanisms for Covid-19 vaccines and participated in discussions with the U.S. Council of Economic Advisors and Domestic Policy Council. The views expressed here are our own. We are grateful for helpful comments provided by Susan Athey, Eric Edmonds, Neil Gandal, Richard Gilbert, Daniel Hosken, Scott Kominers, Ruth Levine, Gerard Roland, the Accelerating Health Technologies team, and seminar participants at the Center for Advanced Study in the Behavioral Sciences at Stanford University, Dartmouth College, M.I.T., University of Washington, the Conference on Development Economics and Market Design at Harvard Business School, the Conference in Honor of Eric Maskin at Harvard University, the IO Theory Conference at M.I.T., the International Industrial Organization Conference session on "Innovation and Technology Adoption in Healthcare," and the Market Shaping Conference at Stanford University. We thank Egor Abramov, Alexandre Simoes Gomes, and Kevin Xie for excellent research assistance. Levin thanks the Center for Advanced Study in the Behavioral Sciences for its hospitality and Snyder the Center for Global Development for generous funding. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Michael Kremer, Jonathan D. Levin, and Christopher M. 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.

Designing Advance Market Commitments for New Vaccines  
Michael Kremer, Jonathan D. Levin, and Christopher M. Snyder  
NBER Working Paper No. 28168  
December 2020  
JEL No. D02,I18,O19,O31

### **ABSTRACT**

Advance market commitments (AMCs) provide a mechanism to stimulate investment by suppliers of products to low-income countries. In an AMC, donors commit to a fund from which a specified subsidy is paid per unit purchased by low-income countries until the fund is exhausted, strengthening suppliers' incentives to invest in research, development, and capacity. Last decade saw the launch of a \$1.5 billion pilot AMC to distribute pneumococcal vaccine to the developing world; in the current pandemic, variations on AMCs are being used to fund Covid-19 vaccines.

This paper undertakes the first formal analysis of AMCs. We construct a model in which an altruistic donor negotiates on behalf of a low-income country with a vaccine supplier after the supplier has sunk investments. We use this model to explain the logic of an AMC—as a solution to a hold-up problem—and to analyze alternative design features under various economic conditions (cost uncertainty, supplier competition). A key finding is that optimal AMC design differs markedly depending on where the product is in its development cycle.

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

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

Jonathan D. Levin  
Graduate School of Business  
Stanford University  
Stanford, CA 94305-7298  
and NBER  
jdlevin@stanford.edu

# 1. Introduction

Mechanisms such as patents and prizes that stimulate research and development (R&D) for products sold in high-income markets may fall short in low-income markets. Patents generate dead-weight loss along with the monopoly rents intended to incentivize investment; furthermore, the monopoly rents may be limited in countries with mostly poor consumers, particularly if the country or aid agency acting on its behalf ignores these patents or uses bargaining power or public pressure to push down prices. Prizes may lead to the development of products that, while meeting the letter of the competition's technical specifications, fail to meet consumers' true needs.

The difficulty in meeting the needs of poor countries is particularly apparent in the market for vaccines. Vaccines are a highly cost-effective tool to improve global public health.<sup>1</sup> Yet the provision of vaccines in poor countries lags widespread use in rich countries and the development of vaccines targeting diseases of poor countries has been disappointingly slow.<sup>2</sup> This situation has sparked a host of initiatives to catalyze vaccine markets in developing countries. Among the most prominent has been an Advance Market Commitment (AMC) piloted last decade for a pneumococcal vaccine.

The idea of using AMCs for vaccines was proposed by Kremer and Glennerster (2004). In advance of firms' investments in R&D, donors pledge a fund to subsidize initial purchases of a newly introduced vaccine above and beyond unit production costs. This feature helps overcome the concern on the part of firms that they will not be able to recoup the large investments required to develop a new vaccine for poor countries, faced with consumers who lack sufficient income to constitute significant demand on their own and aid agencies liable to use their bargaining power to "hold up" the firms' investments, pushing the price toward unit production cost in ex post negotiations. The "market" component of the AMC is that the subsidy is triggered only when the low-income country makes a purchase decision, requiring it to make a small copayment. This feature helps avoid an outcome in which the firm produces a product meeting the technical specifications but not delivering much consumer utility for reasons that are hard to specify in a contract.

Since donors' budget rules typically do not allow indefinite commitments, and since health needs and opportunities are subject to change over time, in practice donors will not agree to subsidize vaccine purchases indefinitely, leaving open the question of market design in the so-called

---

<sup>1</sup>See the cost-effectiveness estimates in Ozawa et al. (2012).

<sup>2</sup>Adoption of vaccines developed in the 1980s, such as Hepatitis B and Hib, on childhood immunization schedules of most rich countries, only began in poor countries decades later. Highly effective vaccines for diseases that are more prevalent in poor countries such as malaria, tuberculosis, yellow fever, and HIV, await development.

“tail period,” after the AMC fund is exhausted. The original AMC design called for firms to offer the vaccine at a lower, sustainable price in the tail period as *quid pro quo* for the subsidy received. This feature helps overcome the concern on the part of poor countries and aid agencies that they adopt a vaccine that in the long run becomes prohibitively priced. In principle, an AMC allows companies to realize an early economic return on their investment while avoiding deadweight loss from monopoly pricing in adopting countries over a longer horizon.

The original idea of an AMC focused on technologically distant projects, with the goal of stimulating the substantial additional R&D required for a marketable product. However, perhaps aiming for a quicker “win” (Cernuschi et al. 2011), the first AMC focused on a technologically closer product. In 2007, the Gates Foundation and a group of countries pledged \$1.5 billion toward a pilot AMC for a next-generation vaccine against pneumococcus, the leading cause of severe pneumonia worldwide, killing over 700,000 children under five each year (World Health Organization 2007). The AMC was intended to incentivize completion of late-stage clinical trials and, importantly, production capacity to serve the low-income market (as much as 210 million additional doses each year to achieve universal vaccination). An Economics Expert Group was presented with an AMC designed according to the specifications of the Framework Document (World Bank and GAVI 2006) and commissioned to specify parameters such as the subsidy rate, price cap during the tail period, and inflation indexing. The Economics Expert Group ended up suggesting modifications to the framework AMC, which were incorporated into the pilot program launched by GAVI (formerly the Global Alliance for Vaccines and Immunizations) in 2009.<sup>3</sup>

Since that pilot, AMCs have grown in relevance. In the Covid-19 pandemic, calls were issued for the launch of an AMC to accelerate development and production of vaccines (Athey et al. 2020). GAVI announced a \$2 billion international funding mechanism, COVAX, containing an component explicitly labeled an AMC, offering manufacturers volume guarantees in advance of licensure to expand supply to developing countries (GAVI 2020). Under Operation Warp Speed, the U.S. government offered contracts of widely varying forms to manufacturers; the contract with Pfizer commits to a \$2 billion payment contingent on licensure or emergency use authorization for 100 million doses (Health and Human Services 2020). Proposed AMCs have not been limited to vaccines, with applications ranging from drug treatments for tuberculosis (Chalkidou et al. 2020b) to low carbon technologies (Department for International Development 2009), to agricultural innovations (von Braun, et al. 2012).

---

<sup>3</sup>See Kremer, Levin, and Snyder (2020) for more detail on the history of the pilot AMC for pneumococcal vaccine.

This paper provides the first formal analysis of AMCs. Besides solidifying the logic behind AMCs, a formal model facilitates the evaluation of alternative program designs and characterization of optimal parameters for those policies. Although we are chiefly interested in making general points about AMCs, we will spend some time focusing on the design of the pneumococcal pilot, examining whether the modifications to the framework design suggested by the Economics Expert Group have a theoretical justification, and whether further modifications might result in further improvements. The theory can help clarify which of the panoply of contracts offered in in the Covid-19 pandemic are appropriately labeled AMCs and which features of those contracts reflect sound design principles.

A broad theme emerging from the analysis is that the optimal AMC design can differ dramatically depending on how far along the product is in its development cycle. For a technologically close product where most of the required R&D investment has been sunk, the goal is to spur investment in capacity to serve low-income countries. We show that subtle changes in AMC design can have large effects on capacity investment. The framework design, which subsidizes initial purchases using a fixed fund, and may do a good job stimulating R&D investment, does not stimulate much additional capacity investment without intense supplier competition. Indeed, it has zero effect on a monopoly's incentives if the fund accrues interest. Better incentives are provided if the real value of the fund decreases over time, better still is the design requires a supply-commitment that matches the AMC fund to committed supply. This result supports the AMC design ultimately used for the pneumococcus AMC, where only two suppliers were initially viable. We show that capacity incentives can be improved further with a capacity-forcing agreement that takes away the firm's discretion over capacity, possibly attaining the first best even for a monopoly supplier.

The design of AMCs for technologically distant products differs in three ways. First, the AMC must induce R&D in addition to capacity investment, calling for larger subsidies. Second, the firm may have less private information about production costs early in the product's development, which improves AMC efficiency because, from a mechanism-design perspective, less private information means a smaller information rent has to be paid to the firm. Third, a complete contract over the product's technical specifications is difficult to write far in advance of product development, reinforcing the role of a country copayment as a sort of "kill switch" for products meeting the specifications but providing little consumer utility for numerous reasons that are hard to anticipate.

We develop these results in a simple model with three players: a firm, a poor country, and a donor which makes vaccine purchases on behalf of the country. The donor specifies the terms

of the AMC or other commitment contract. The firm (initially a monopoly for simplicity) makes investments, which can include investments in R&D in the case of a technologically distant product or just in capacity in the case of a technologically close product. This leads to an ex post period during which the donor and firm bargain over vaccine purchases. The absence of an AMC or other ex ante policy does not rule out the supply of the vaccine ex post, but capacity is inefficiently low. Similarly, the existence of an AMC does not preclude subsequent bargaining over price. So the design of an AMC must anticipate the bargaining game that occurs when the vaccine reaches the market.

Regarding related literature, the idea of using an AMC to fund vaccine purchases was proposed by Kremer and Glennerster (2004) building on ideas in Kremer (2000a, 2000b). Further refinements were provided by Levine, Kremer, and Albright (2005); Berndt and Hurvitz (2005); and Kremer and Williams (2010). These papers were largely conceptual; the present paper contributes a formal theoretical analysis. Kremer, Levin, and Snyder (2020) discusses practical lessons from the pneumococcus AMC and evidence of its effectiveness. Berndt et al. (2007) provide empirical forecasts of costs and effectiveness of AMCs for neglected diseases. Snyder, Begor, and Berndt (2011) provides an early assessment of the pneumococcal pilot and calibrations of alternative policy designs. In the current Covid-19 pandemic, several articles have laid out the conceptual case for a vaccine AMC (e.g., Chalkidou et al. 2020a); Snyder et al. (2020) calibrate the optimal size of such a program. AMCs have been the subject of critiques ranging from high program costs (Light 2005, Scudellari 2010) to the ethical dilemma raised by respecting developing countries' preferences for culturally acceptable but less effective vaccines (Sonderholm 2011).

Our paper also contributes to a wider literature on innovation incentives. A few related contributions include Acemoglu and Lin (2004) on the contribution of market size to innovations in pharmaceutical markets; Kremer (1998) on adding patent buyouts to the intellectual-property toolkit; Ridley, Grabowski, and Moe (2006) on allocating of priority review vouchers as an inducement to invest in pharmaceuticals for neglected tropical diseases; and Weyl and Tirole (2012) on mechanism that is a hybrid of patents and prizes, allowing a continuous choice between the two, optimally trading off rents paid to firms versus deadweight loss.

The plan of the paper is as follows. Section 2 presents the model. Section 3 begins the analysis with the basic case of a monopoly firm producing a technologically close product. The section examines a series of increasingly efficient policies, from a framework AMC to a supply commitment to a forcing contract. The section ends by cataloging a variety of refinements that might be added

to these basic policies such as price caps, purchase guarantees, escrowed interest, procurement through an agent, and country copayments. While analysis of these refinements does contribute to a conceptual understanding of AMCs, the main contribution is to raise considerations of practical policy interest; so for space considerations we relegate the details of the analysis to the appendix. Section 4 extends the analysis to allow for asymmetric information about the firm's costs. Section 5 extends the analysis from a monopoly to allow for an arbitrary number of firms. Section 6 moves the focus from a technologically close product for which R&D costs have not yet been sunk. The last section concludes. Appendix A provides proofs omitted from the text. Appendix B provides a full analysis of the refinements of the basic AMC cataloged at the end of Section 3.

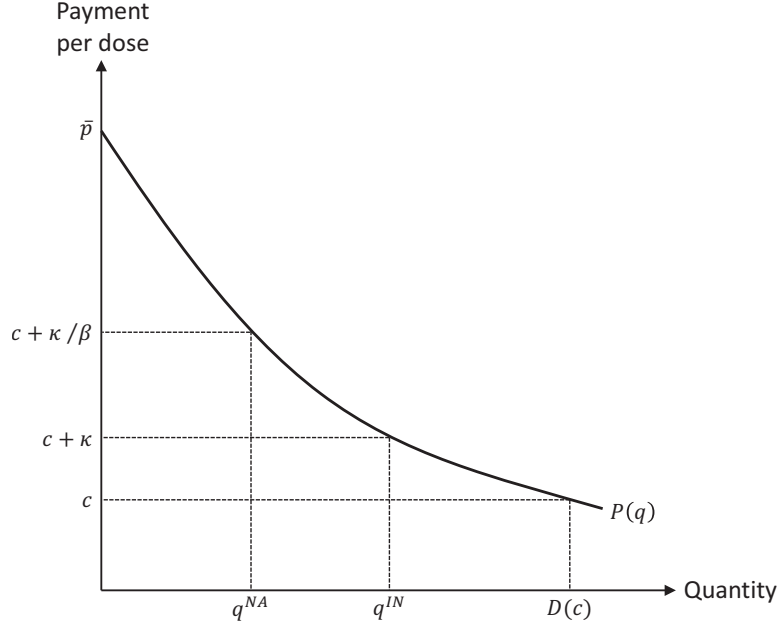
## 2. Model

The model has two periods, *ex ante* and *ex post*, the dividing line being the point at which the firm sinks any required investments. The *ex post* period involves continuous time indexed by  $t$ . Let  $r > 0$  be the market interest rate as well as all players' internal discount rates.

The model has three players: a firm, country, and donor. A profit-maximizing firm has the opportunity to develop a new product that it can supply to a representative low-income country. The analysis initially focuses on a monopoly firm, extended later to imperfect competition. For concreteness we will take the product to be a vaccine, though it could represent any new good or service. The country obtains health benefit  $B(q)$  from the vaccination of  $q$  of its citizens. This flow benefit can be realized each  $t$  because of new cohorts of unvaccinated newborns. Normalize  $B(0) = 0$ . Assume that  $B'(q) > 0$  and  $B''(q) < 0$  for all  $q \geq 0$  and that  $\lim_{q \rightarrow \infty} B'(q) = 0$ .

Initially suppose that the country cannot pay for the vaccine itself, either because it is too poor to afford any payment or because its leaders, who may or may not be benevolent, have other priorities. (Section 6.1 and Appendix B extend the model to allow for country contributions.) A donor, who could represent one or more non-governmental organizations and/or high-income countries with development-aid budgets, has regard for the low-income country's welfare. Its flow utility is  $\alpha B(q) - X$ , where  $\alpha \in (0, 1)$  represents the degree of donor altruism and  $X \geq 0$  is its expenditure on vaccines each instant  $t$ . Let  $P(q) = \alpha B'(q)$  denote the donor's marginal willingness to pay for the vaccine, in effect its inverse demand curve, illustrated in Figure 1.

Establishing some properties of the demand curve streamlines the analysis in later sections. The strict concavity of  $B(q)$  implies  $P'(q) = \alpha B''(q) < 0$  for all  $q \geq 0$ . Hence inverse demand slopes



**Figure 1:** Donor inverse demand

down. Let  $\bar{p} = P(0)$  denote the choke price. That  $B(q)$  is twice continuously differentiable at  $q = 0$  implies  $P'(0) = \alpha B''(0)$  exists, implying  $P(0)$  is finite, and thus  $\bar{p}$  is finite. Let  $\tilde{D}(p) = P^{-1}(p)$  denote the inverse of  $P(q)$ . The following lemma establishes useful properties of this inverse.

**Lemma 1.** *For all  $p \in (0, \bar{p})$ ,  $\tilde{D}(p)$  is a finite, positive number and  $\tilde{D}'(p) < 0$ .*

The proof, relying on the inverse function theorem, is provided in the appendix. Regarding the endpoints of the interval, it is easy to see that  $\lim_{p \rightarrow 0} \tilde{D}(p) = \infty$  and  $\tilde{D}(\bar{p}) = 0$ . Define the demand curve  $D(p)$  to be the following extension of  $\tilde{D}(p)$  to all positive prices, i.e.,  $D(p) = \tilde{D}(p)$  for  $p \in (0, \bar{p})$  and  $\tilde{D}(p) = 0$  for  $p \geq \bar{p}$ . The properties of  $\tilde{D}(p)$  established in the lemma are inherited by  $D(p)$  for all  $p \in (0, \bar{p})$ .

The firm has three sources of cost. Let  $R$  be the research and development cost that must be sunk at time 0 for the product to exist. Let  $KQ$  be the cost that must be sunk at time 0 to install capacity  $Q$ , where  $K > 0$ . Notice that  $R$  is a fixed cost, independent of capacity or output, but  $KQ$  is increasing in capacity (linearly for simplicity). It will be convenient to express these sunk costs, which are stocks, as flows to facilitate comparison to other flows in the model. To this end, let  $\rho = rR$  and  $\kappa = rK$  be the per-period payments to finance the research and development and unit capacity investments, respectively. Once sunk investments are in place, the cost of producing at rate  $q \leq Q$  each period is  $cq$ , where  $c > 0$  is the marginal cost of producing a unit in a period.



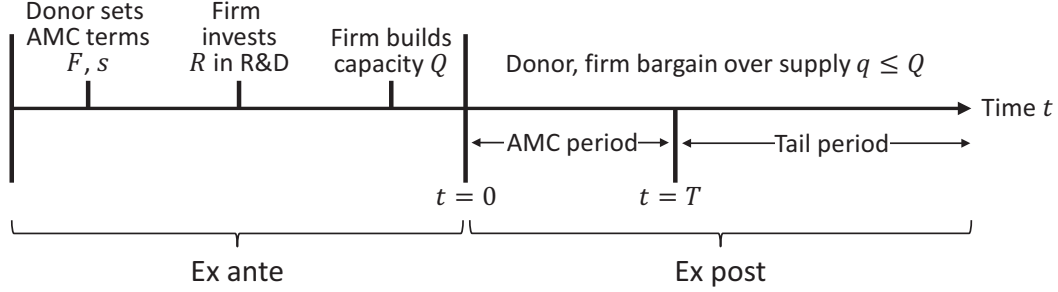
We will analyze a series of AMC designs, but to fix ideas we describe the framework AMC. In the framework AMC, the donor sets a fund size  $F$  and a per-unit subsidy  $s$ . Imparting commitment power to the AMC, the fund is assumed to be locked in an escrow that cannot be used for other purposes.<sup>4</sup> As a baseline, suppose that interest earned by the escrow accrues there and augments the committed funds. (We also will analyze the variant in which interest earned by the escrow flows back to the donor.) The firm receives the subsidy  $s$  for each unit purchased by the country until the fund is exhausted. Let  $T$  denote the time at which this happens. Time interval  $(0, T)$  is the AMC period and  $(T, \infty)$  the tail period. The framework AMC is linear in output. Nonlinear contracts (e.g., forcing contracts) are also possible, as are contracts that are functions of variables other than output (e.g, capacity). Our approach will be to analyze various contractual forms and discuss the assumptions on contractual completeness and/or commitment power required for the form to be feasible in the relevant passages below.

There are a variety of ways to model the process of price formation and purchasing ex post. We assume that once the firm sinks its investments, it engages in Nash bargaining with the donor over the sequence of future vaccine purchases. A hold-up problem arises from the fact that bargaining takes place after investments have been sunk. Using a bargaining framework to model vaccine purchases has several virtues. Besides hold up, it builds in elements of bilateral monopoly, both important and realistic features of markets in which AMCs might be used. The absence of an AMC does not preclude the possibility of trade, nor does its presence preclude the possibility of bargaining under its shadow. Thus, while the AMC will have scope to affect equilibrium outcomes, it will not be through exogenous assumptions on the effect of the AMC on the structure of price formation. Let  $\beta \in (0, 1)$  be the firm's bargaining weight and  $1 - \beta$  the donor's.

A complete formulation of Nash bargaining requires several additional details to be delineated. We assume that parties immediately engage in a once-for-all Nash bargain at the start of the ex post period, at  $t = 0$ , covering the entire sequence of vaccine prices and quantities each instant  $t$  of continuous time, to which they can commit. The outcome is equivalent to the Markov-perfect equilibrium of instant-by-instant Nash bargaining in the present setting of a stationary game with

---

<sup>4</sup>Looking inside the black box of the escrow, the assumption that the escrow can be committed to a single purpose can be viewed as prohibiting the escrow administrator from bilaterally renegotiating with the donor ex post over the use of the funds. Allowing the firm to bring a legal action against such a reallocation would restore the commitment value of the escrow. If the escrow funds are not committed to a single purpose and instead are fungible for the donor, one can show that the framework design of the AMC provides no investment incentives. As discussed in Appendix B, having an escrow in the model under some circumstances is equivalent to having a model in which the donor is split into two players, a principal who donates the funds and designs the AMC and an agent who carries out the ex post bargaining. The latter model resembles the structure of the pilot AMC for pneumococcus.



**Figure 2:** Timeline for model

efficiency bargaining under perfect certainty with freely transferable surplus, but is mathematically more tractable. Another detail concerns parties' threat points. Parties' threat points depend on the policy in force at  $t = 0$ . In the absence of an AMC or other commitment, we assume trade requires the active assent of both parties, implying that parties obtain zero surplus in the continuation game following a breakdown of bargaining. The threat point when an AMC is in place is not well-defined without further specification. We assume as a baseline that AMCs (or any other policy imposed ex ante) require the active assent of both bargaining parties for trade to occur, just as in the absence of an AMC. A later section will analyze the alternative in which the AMC empowers the firm to trade at the specified terms without the assent of the donor, labeled a purchase guarantee.

To summarize the model's timing, depicted in Figure 2, the ex ante period begins with the donor setting the terms of the AMC. Observing this, the firm sinks any required investment in research, development, and/or capacity. This leads to the ex post period, starting at time 0. The donor and firm engage in Nash bargaining over the supply of  $q$  units of vaccine to the country up to capacity  $Q$ . The once-for-all bargain at time  $t = 0$  covers the sequence of vaccine prices and quantities for all  $t$  thereafter.

To rule out trivial cases, we maintain the following condition throughout the analysis.

**Assumption 1.**  $c + \kappa < \bar{p}$ .

We will see in the analysis of the integrated benchmark, in which the donor is allowed to control the operations of the firm directly, that optimal capacity and output is  $D(c + \kappa)$ . Assumption 1 combined with Lemma 1 implies  $D(c + \kappa) > 0$ . Hence the integrated donor would install some capacity. If the assumption did not hold, then not only the integrated benchmark but all equilibria in which the donor bargains with a separate firm would involve the trivial outcome of no capacity or supply.

### 3. Technologically Close Products

We begin the analysis with the simple case of a near-term product, specifically a vaccine that is far enough in research and development before the AMC is implemented that the only substantial investment required is the expenditure  $KQ$  for capacity, not R&D expenditure  $R$ . Assume  $K > 0$ , implying  $\kappa > 0$ . The product is also far enough along that its attributes can be specified fairly completely. For simplicity, we will take the extreme case in which these attributes can be specified perfectly, implying that ex ante commitments can specify ex post payments conditional on  $q$  (in short,  $q$  is contractible).

To facilitate the analysis, before analyzing different AMC designs, we study two benchmarks at the extremes of the range of possibilities. We first analyze the extreme with the greatest incentive for capacity investment: a counterfactual integrated outcome in which the donor directly controls the firm's operations. We then analyze the extreme with the least incentive for capacity investment: the outcome in the absence of integration and in the absence of an AMC. The outcome with an AMC will fall somewhere between these two benchmarks. We start our analysis of AMCs in Section 3.4 with the framework design, and then analyze possible improvements moving to supply commitments and forcing contracts.

#### 3.1. Integrated Benchmark

Suppose the donor can directly control the operations of the firm. Let  $Q^{IN}$  be the donor's optimal capacity and  $q^{IN}$  flow quantity in this benchmark, maximizing the present discounted value of benefits minus costs,

$$\int_0^{\infty} [\alpha B(q) - cq] e^{-rt} dt - KQ = \frac{1}{r} [\alpha B(q) - cq - \kappa Q], \quad (1)$$

subject to

$$q \leq Q. \quad (2)$$

Since expending joint resources on idle capacity is inefficient, the constraint obviously binds. Substituting the constraint treated as an equality into (1) yields

$$\frac{1}{r} [\alpha B(Q) - (c + \kappa)Q] \equiv w(Q, \theta). \quad (3)$$

Since the expression for joint welfare on the left-hand side reappears frequently, we streamline the subsequent analysis by introducing the shorthand  $w(Q, \theta)$  for it, where  $\theta = (r, \alpha, c, \kappa, \beta)$  is the vector of model parameters.

The first-order condition for maximizing (3) with respect to  $Q$  is  $P(Q) - (c + \kappa) = 0$ , implying  $Q^{IN} = q^{IN} = D(c + \kappa)$ , which is positive under Assumption 1 by Lemma 1. The following proposition summarizes this analysis.

**Proposition 1.** *Capacity and output for a monopoly firm in the integrated benchmark are given by  $Q^{IN} = q^{IN} = D(c + \kappa) > 0$ .*

### 3.2. Benchmark Without AMC

We next analyze the outcome in the absence of integration or an AMC. Backward induction starting from the ex post Nash bargain, then folding the game back to the firm's capacity choice in the ex ante period, yields the following objective function:

$$\beta \int_0^\infty [\alpha B(q) - cq] e^{-rt} dt - KQ = \frac{\beta}{r} \left[ \alpha B(Q) - \left( c + \frac{\kappa}{\beta} \right) Q \right] \quad (4)$$

$$\equiv \pi(Q, \theta). \quad (5)$$

To see the expression on the left-hand side of (4), since either party can veto trade, their threat points are both 0. The firm's Nash bargaining surplus is the share  $\beta$  of the gains from trade given by the integral. Subtracting capacity cost  $KQ$  yields the firm's ex ante profit function. The right-hand side of (4) follows from the fact that a positive capacity cost and efficient ex post bargaining leads all capacity to be utilized each instant in equilibrium. Substituting  $q = Q$  in the left-hand side of (4) and rearranging gives the right-hand side. Since the right-hand side of (4), reflecting the firm's ex ante profit in the absence of an AMC, reappears frequently, we streamline the analysis by introducing the notation  $\pi(Q, \theta)$  in (5), where  $\theta$  is the vector of model parameters.

The first-order condition for maximizing (4) with respect to  $Q$  is  $P(Q) - (c + \kappa/\beta) = 0$ , implying  $Q^{NA} = D(c + \kappa/\beta)$ . Thus,  $Q^{NA} = D(c + \kappa/\beta) < D(c + \kappa) = Q^{IN}$ , where the inequality follows from the fact that demand is decreasing in the relevant range by Lemma 1. We have thus sketched the following proposition. The formal proof in the appendix fills in missing details including that all capacity is utilized for output in equilibrium.

**Proposition 2.** *Capacity and output for a monopoly firm in the benchmark with no AMC are given by  $Q^{NA} = q^{NA} = D(c + \kappa/\beta)$  and the ex ante present discounted value of the firm's profits by  $\pi(Q^{NA}, \theta)$ . Comparing capacity to that in the integrated benchmark, we have  $Q^{NA} < Q^{IN}$ ,  $\lim_{\beta \rightarrow 0} Q^{NA} = 0$ , and  $\lim_{\beta \rightarrow 1} Q^{NA} = Q^{IN}$ .*

That  $Q^{NA} < Q^{IN}$  is a key result of the paper, embodying the hold-up problem which AMCs are designed to address. The hold-up problem arises because returns from the firm's sunk investment in capacity are appropriated by the donor in ex post negotiations unless the firm has 100% of the bargaining power. In the absence of an AMC, this leads the firm to underinvest in capacity compared to the integrated benchmark. In the limit  $\beta \rightarrow 0$  in which the firm has no bargaining power, the hold-up problem is so severe that the firm installs no capacity. In the limit  $\beta \rightarrow 1$  in which the firm has all the bargaining power, the hold-up problem disappears.

Figure 1 illustrates equilibrium quantity in the two benchmarks. Integrated quantity  $D(c + \kappa)$  reflects marginal-cost pricing, taking a comprehensive notion of marginal cost, including both the unit cost of production  $c$  and that unit's share  $\kappa$  of the cost of capacity. In the no-AMC benchmark, the quantity  $D(c + \kappa/\beta)$  gives extra weight to that part of the cost  $\kappa$  subject to hold up, where the extra weight is inversely proportional to the firm's bargaining power.

### 3.3. Principle of Full Capacity Utilization

Neither benchmark analyzed in the previous two subsections involved excess capacity. The key to this result is that both are ex post efficient, a consequence of integration in one case and Nash bargaining in the other. The ex post efficient outcome is for all capacity  $Q$  being used as long as the marginal benefit  $\alpha B'(Q)$  exceeds the marginal production cost  $c$ , or rearranging, as long as  $Q \leq D(c)$ . But  $D(c)$  is more capacity than even an integrated firm would install facing any positive capacity cost  $\kappa$ . Hence all capacity is used each instant.

This insight will apply more generally to all the equilibria in the remainder of this section. A positive capacity cost will induce the firm to install less capacity than  $D(c)$ ; efficient ex post bargaining will then lead to all capacity being utilized each instant. We will invoke this result in the remainder of the section using the shorthand *principle of full capacity utilization*.

Deriving the principle for the integrated and no-AMC benchmarks was easy because changes in the quantity traded in any instant had no external effects on the surplus to be shared in future instants. A caveat in applying the principle to the case of a AMC is that producing more in the current instant draws the fund down, reducing the surplus that can be shared later. The principle

applies, but some work is required in the proofs to make the argument.

### 3.4. Framework AMC

The first AMC design we will analyze is the framework design, involving a fund  $F$  placed in escrow ex ante from which a subsidy  $s$  is paid for each unit of the vaccine purchased ex post until the fund is exhausted. Start by positing that the duration of the AMC period  $T$  is finite. Then  $T$  can be determined by the following accounting identity:

$$F = \int_0^T sqe^{-rt} dt = \frac{sq}{r} (1 - e^{-rT}). \quad (6)$$

The left-hand side of (6) is the lump sum paid into the escrow ex ante. The right-hand side is the present discounted value of the flow of subsidy payments  $sq$  made each instant during the AMC period. The discounting reflects the baseline specification for AMC design that interest on funds accrues to the escrow.

We again solve for the equilibrium using backward induction starting from Nash bargaining at  $t = 0$ . The model requires both parties to assent to trade during the AMC period and tail period, implying that their threat points are again 0. Gains from trade equal the present discounted value of joint surplus over the continuation game

$$\int_0^\infty \alpha B(Q) e^{-rt} dt + \int_0^T (s-c) Q e^{-rt} dt - \int_T^\infty c Q e^{-rt} dt = \frac{1}{r} [\alpha B(Q) - cQ + sQ(1 - e^{-rT})]. \quad (7)$$

The first term on the left-hand side is the donor's gross surplus (not including bargaining payments) over the whole ex post period, the second term is the firm's gross surplus (again not including bargaining payments) during the AMC period, and the third is the firm's gross surplus during the tail period. The right-hand side of (7) follows from integrating and rearranging. Several additional notes about (7) are in order. Note that  $s$  appears in the second term. While one might think it should net out of joint surplus as a mere transfer, because it was sunk in a dedicated escrow, ex post subsidy payments are not subtracted from the donor's contribution to joint surplus. Note further that capacity  $Q$  has been substituted for output in (7) on the basis of the principle of full capacity utilization. This principle obviously holds in the tail period because reducing output below  $Q$  has no later benefits. Work is required to verify that the principle holds during the AMC period because a reduction in output in one instant leaves more funds in the escrow, providing the future benefit

of extending the AMC period. We relegate this verification to the proof of Proposition 3 in the appendix. For now, we just note that it is intuitive that the principle holds during the AMC period. Deferring the draw down of AMC funds by reducing output has a neutral effect on joint surplus because the funds accrue interest at the same rate in or out of the escrow. The only remaining effect of an output reduction is the inefficiency from producing less than the capacity-constrained optimum,  $Q$ .

Fold the game back to the firm's ex ante choice of capacity. Given its threat point is 0, the its continuation payoff from Nash bargaining is its share  $\beta$  of the gains in (7). Subtracting capacity costs yields the following expression for ex ante profit,

$$\frac{\beta}{r}[\alpha B(Q) - cQ + sQ(1 - e^{-rT})] - KQ = \pi(Q, \theta) + \beta F. \quad (8)$$

The right-hand side follows from substituting from (6) in (8) and then using (5).

Rearranging the first-order condition from maximizing (8) with respect to  $Q$  yields  $Q^{TS} = D(c + \kappa/\beta) = Q^{NA}$ . We have obtained a stark result: capacity and output are the same with a framework AMC as without; the donor receives nothing for the funds  $F$  it contributes to the AMC. Mathematically, the contribution of the AMC to the firm's surplus, which is embodied in equation (8) by the term  $\beta F$ , is not a function of  $Q$ . Intuitively, if the firm reduces its capacity, it draws down the AMC at a slower rate, but this leaves more funds in the escrow account accruing interest at the same rate as the firm's own bank account. The firm ends up being indifferent as how fast it draws down the AMC fund. The AMC thus contributes nothing to incentives to invest in capacity.

**Proposition 3.** *Consider any terms  $s > 0$  and  $F > 0$  for which the framework AMC has temporary subsidy. The framework AMC adds no incentives for a monopoly firm to invest in capacity:  $Q^{TS} = q^{TS} = D(c + \kappa/\beta) = q^{NA} = Q^{NA}$ .*

### 3.5. Perpetual Subsidy

The donor would never make a positive contribution  $F$  toward this useless program. The result that an AMC with a temporary subsidy is useless leaves open the possibility that an AMC with a perpetual subsidy may be useful. With a temporary subsidy, an increase in output in instant  $t$  merely shifts the timing of the subsidy payouts with no net effect on the total earned from the program and thus no effect on incentives. With a perpetual AMC, a current increase in output does not reduce the subsidy that can be earned in the future—promised to be  $s$  per unit regardless of

how much was produced before. Hence a perpetual subsidy may enhance investment incentives.

A perpetual subsidy requires a sufficiently large endowment  $F$  given  $s$  and  $q$ , namely  $rF \geq sq$ . The donor prefers the smallest fund satisfying that condition:

$$F = \frac{sq}{r}. \quad (9)$$

We are left to solve for the optimal  $s$ . We again use backward induction, starting from the ex post Nash bargain at  $t = 0$  over vaccine supply, positing some arbitrary  $s$  set by the donor.

For Nash bargaining to lead to trade, both parties must assent, implying that their threat points are 0 as before. Gains from trade equal the present discounted value of joint surplus over the continuation game

$$\int_0^{\infty} [\alpha B(Q) - cQ + sQ] e^{-rt} dt = \frac{1}{r} [\alpha B(Q) - cQ + sQ]. \quad (10)$$

We have substituted capacity  $Q$  for quantity, invoking the principle of full capacity utilization, which obviously applies in this case. As was the case above with a temporary subsidy, here, too,  $s$  appears in the expression for continuation surplus because the subsidy is paid out of a dedicated escrow, the funds in which do not have an opportunity cost to the donor.

Folding the game back to the ex ante period, the firm's equilibrium capacity choice maximizes its share  $\beta$  of (10) less capacity costs, which upon rearranging becomes

$$\frac{\beta}{r} \left[ \alpha B(Q) - \left( c + \frac{\kappa}{\beta} \right) Q + sQ \right] = \pi(Q, \theta) + \frac{\beta s Q}{r}. \quad (11)$$

The first-order condition from maximizing (11) with respect to  $Q$  is

$$P(Q) - \left( c + \frac{\kappa}{\beta} \right) + s = 0. \quad (12)$$

This is the same as the first-order condition from the benchmark with no AMC except that  $s$  has been added, enhancing incentives when  $s > 0$ .

Folding the game back, the donor chooses the AMC terms to maximize the present discounted value of its  $1 - \beta$  share of the joint surplus (10) from Nash bargaining less the fund  $F$  endowing the



perpetual subsidy:

$$\frac{1-\beta}{r}[\alpha B(Q) - cQ + sQ] - F = \frac{1}{r} \{ (1-\beta)[\alpha B(Q) - cQ] - \beta sQ \}. \quad (13)$$

The right-hand side of (13) follows from substituting for  $F$  from (9) and rearranging. The donor's problem is to choose  $s$  to maximize (13), where  $Q$  is implicitly a function of  $s$  via (12), which is in essence an incentive-compatibility constraint.

The one-to-one correspondence between capacity and the subsidy implicit in (12) allows us to express the donor's problem equivalently in terms of choice variable  $Q$  rather than  $s$ , yielding considerable insight. Solving (12) for  $s$  in terms of  $Q$ , substituting into (13), and rearranging gives the following expression for the donor's objective function:

$$\frac{1}{r}[(1-\beta)\alpha B(Q) + \beta QP'(Q) - (c+\kappa)Q]. \quad (14)$$

Taking the first-order condition from maximizing (14) with respect to  $Q$  and rearranging,

$$P(Q) + \beta QP'(Q) - (c+\kappa) = 0. \quad (15)$$

Rearranging further yields the following implicit solution for equilibrium capacity under a perpetual subsidy, denoted  $Q^{PS}$ :

$$Q^{PS} = D(c+\kappa - \beta Q^{PS}P'(Q^{PS})). \quad (16)$$

Because  $-\beta P'(Q^{PS})Q^{PS} > 0$  for  $\beta > 0$ ,  $D(c+\kappa - \beta P'(Q^{PS})Q^{PS}) < D(c+\kappa) = Q^{IN}$ . Hence, if the firm has any bargaining power, the perpetual subsidy cannot achieve the integrated outcome. We have proved the following proposition.

**Proposition 4.** *The framework AMC that is optimal for the donor involves a perpetual subsidy,  $s^{PS} = c + \kappa / \beta - P(Q^{PS})$ , where  $Q^{PS}$  is the monopoly firm's equilibrium capacity under this policy given by the implicit solution to (16). For all  $\beta > 0$ ,  $Q^{PS} < Q^{IN}$ , but  $\lim_{\beta \rightarrow 0} Q^{PS} = Q^{IN}$ .*

Further insight can be obtained by rearranging equation (16) in the form of the Lerner index:

$$L^{PS} = \frac{P(Q^{PS}) - (c+\kappa)}{P(Q^{PS})} = -\beta \frac{Q^{PS}P'(Q^{PS})}{P(Q^{PS})} = \frac{\beta}{|\eta^{PS}|}. \quad (17)$$

The first equality reflects the definition of the Lerner index as the percentage markup of price over marginal cost, but uses a comprehensive notion of marginal cost appropriate to this context,

including production and capacity costs. The second equality follows from rearranging (16). The third equality follows from defining  $\eta^{PS}$  to be the elasticity of the donor's vaccine demand at the equilibrium quantity. In the limit  $\beta \rightarrow 0$ , the equilibrium converges to efficient, i.e., marginal-cost, pricing. The donor sets  $s$  efficiently ex ante because it appropriates all surplus ex post. As  $\beta$  increases toward 1, the donor increasingly behaves as a standard monopsonist, setting the linear price  $s$  per unit of capacity according to an inverse-elasticity rule. As will be seen momentarily, there is a limit to how close the donor comes to implementing an inverse-elasticity rule because a framework AMC does not exist for  $\beta$  sufficiently close to 1.

The subsection concludes with a more technical discussion of existence of a framework AMC. Call an AMC *nontrivial* if it strictly improves the donor's surplus relative to the equilibrium in the absence of an AMC. Call an AMC *feasible* if the firm would be willing to sign it or in other words it satisfies the firm's participation constraint. We have the following proposition, proved in the appendix.

**Proposition 5.** *There exists a nontrivial, feasible framework AMC (NFFA) if and only if  $Q^{PS} > Q^{NA}$ . To characterize the parameters satisfying this condition, there exists an NFFA for  $\beta$  sufficiently close to 0 and not for  $\beta$  sufficiently close to 1. No NFFA exists for  $\kappa$  sufficiently close to 0.*

*Suppose donor objective function (14) is concave in  $Q$ . Then there exists  $\hat{\beta} \in (0, 1)$  such that an NFFA exists for all  $\beta < \hat{\beta}$  and not for all  $\beta \geq \hat{\beta}$ . Furthermore, there exists  $\hat{\kappa} > 0$  such that an NFFA exists for all  $\kappa \in (\hat{\kappa}, \bar{p} - c)$  and not for  $\kappa \leq \hat{\kappa}$ .*

To gain some intuition for the conditions guaranteeing no NFFA exists, for  $\beta$  close to 1, the firm obtains most of the bargaining surplus so chooses a capacity close to the integrated solution  $Q^{IN}$  in the absence of an AMC. With little need to further enhance incentives with a positive subsidy, the donor would prefer to extract rent with a negative subsidy were this feasible. Similarly, for  $\kappa$  close to 0, the scale of the hold-up problem is limited because of the small scale of ex ante investment. Again the donor would prefer to extract rent with a negative subsidy if this were feasible.

### 3.6. Supply Commitment

The previous subsection showed that a framework AMC can improve investment incentives. To do so, however, required the donor to commit to a perpetuity, which may strain its commitment power in practice. Furthermore, while the AMC with a perpetual subsidy was able to improve incentives, it was not able to achieve the integrated outcome. In this section we analyze an alternative policy to the framework AMC. The Economics Expert Group for the pneumococcus AMC

recommended a move from the framework design to a supply commitment. We will show that the supply commitment can achieve the same outcome as the optimal framework AMC without a perpetual commitment.

In addition to the contractual terms  $F$  and  $s$ , a supply commitment sets a target supply  $\hat{q}$ . Rather than being able to appropriate the entire fund  $F$  given enough time, now the firm is only eligible to appropriate at most a fraction equal to the fraction  $q/\hat{q}$  that it supplies of the target. With this design, the accounting identity determining the length of the AMC period becomes

$$\frac{\min(\hat{q}, q)}{\hat{q}} F = \frac{sq}{r} (1 - e^{-rT}). \quad (18)$$

The left-hand side reflects fund allowed to be used for the subsidy; the right-hand side is the present discounted value of payments out of the fund. An implicit assumption behind the accounting identity is that the duration of the AMC period associated with the supply commitment is finite. We posit this because the presumed benefit of a supply commitment is avoid donor's having to commit to a perpetuity. The min operator in (18) prohibits the firm from collecting more than  $F$  in AMC subsidy payments by producing more than  $\hat{q}$ .

Rearranging (18),

$$F = \frac{s}{r} \max(\hat{q}, q) (1 - e^{-rT}). \quad (19)$$

One can show that if the donor sets  $\hat{q}$  less than the equilibrium output, the supply commitment is identical to a framework AMC. The firm is already collecting the full  $F$ , so further increases in output do not have the beneficial incentive effect of expanding the fund. We are interested in improving on the efficiency of a framework AMC with a supply commitment. Thus we posit that the donor sets the target weakly above equilibrium output:

$$q \leq \hat{q}. \quad (20)$$

We will show that the optimal supply commitment does no worse than the optimal framework AMC, so assumption (20) is without loss of generality. Substituting (20) into (19) and rearranging to solve for the duration of the supply commitment,

$$T = \frac{1}{r} \ln \left( \frac{s\hat{q}}{s\hat{q} - rF} \right). \quad (21)$$

This equation makes clear why a supply commitment improves investment incentives relative to a

framework AMC. With the framework design, an increase in supply increases the subsidy earned each instant during the AMC period, but this is offset by shortening the AMC's duration. With a supply commitment, as (21) shows, any increase in supply up to the target has no effect on the AMC's duration because the fund that the firm is able to draw upon grows proportionately with supply up to the target. This may lead to more powerful investment incentives with a supply commitment than a framework AMC.

Proceeding with the analysis, the derivation of the present discounted value of the firm's ex ante profit proceeds the same here as with the framework design, again giving equation (8). As before we invoke the principle of full capacity utilization to justify substituting capacity  $Q$  for output  $q$ . The only difference is what fills in for  $T$ . Substituting for  $T$  from (21) into the left-hand side of (8) and rearranging yields firm profit

$$\frac{\beta}{r} \left[ \alpha B(Q) - \left( c + \frac{\kappa}{\beta} \right) Q \right] + \frac{\beta Q F}{\hat{q}} = \pi(Q, \theta) + \frac{\beta Q F}{\hat{q}}. \quad (22)$$

Differentiating with respect to  $Q$ , taking  $\hat{q}$  to be exogenous to the firm, and rearranging yields first-order condition

$$P(Q) - \left( c + \frac{\kappa}{\beta} \right) + \frac{rF}{\hat{q}} = 0. \quad (23)$$

Folding the game back to the donor's design of the optimal supply commitment, the present discounted value of donor surplus equals its share  $1 - \beta$  of joint continuation surplus with a finite AMC period—the same as derived in (7)—less its initial contribution  $F$  to the escrow:

$$\frac{1 - \beta}{r} [\alpha B(Q) - cQ + sQ(1 - e^{-rT})] - F = \frac{1 - \beta}{r} \left[ \alpha B(Q) - cQ + \frac{rFQ}{\hat{q}} \right] - F, \quad (24)$$

where the right-hand side follows from substituting for  $T$  from (21) and rearranging. The optimal supply commitment for the donor is given by the values of  $Q$ ,  $F$ , and  $\hat{q}$  maximizing (24) subject to incentive-compatibility constraint (23). We also continue to impose the constraint (20) posited above on target output, which after invoking the principle of full capacity utilization becomes

$$Q \leq \hat{q}. \quad (25)$$

The maximization problem can be simplified by solving (23) for  $F$  and substituting the resulting expression for  $F$  in (24), turning a maximization problem with three choice variables and two constraints into one with two choice variables and one constraint (25). After rearranging, the

objective function from the transformed problem is

$$\frac{1}{r} \left\{ (1-\beta)\alpha B(Q) + \beta QP(Q) - (c+\kappa)Q + (Q-\hat{q}) \left[ c + \frac{\kappa}{\beta} - P(Q) \right] \right\}. \quad (26)$$

It is obvious that (25) binds. Substituting (25) treated as an equality into (26) causes the last term to disappear. Differentiating the resulting objective function with respect to  $Q$  and rearranging yields a first-order condition for the equilibrium capacity identical to (16) from the optimal framework AMC with a perpetual subsidy, leading immediately to the following proposition.

**Proposition 6.** *The donor's surplus from the optimal supply commitment is the same as from the optimal framework AMC with a perpetual subsidy. The required funds are the same across the two policies ( $F^{SC} = F^{PS}$ ) as are equilibrium capacities and outputs ( $Q^{SC} = q^{SC} = q^{PS} = Q^{PS}$ ).*

The proof in the appendix fills in the remaining detail about equality between fund sizes.

Turn next to the question of when a feasible, nontrivial supply commitment exists. A feasible supply commitment must satisfy the firm's participation constraint, which here can be shown to be equivalent to a non-negative subsidy rate  $s \geq 0$ , or equivalently a non-negative fund  $F \geq 0$ . Given that a supply commitment attains the outcome under a framework AMC with a perpetual subsidy, not surprisingly the conditions under which a feasible, nontrivial supply commitment exists are identical to those provided in Proposition 5 for a feasible, nontrivial framework AMC.

Proposition 6 is silent about the subsidy rate  $s^{SC}$  in the optimal supply commitment. That is because this rate is not pinned down in an optimum. A low rate paid over a long AMC period or a high rate paid over a short period can both be optimal. Specifically, any pairs  $(s, T)$  satisfying  $s > 0$ ,  $T > 0$ , and

$$\frac{s}{r}(1 - e^{-rT}) = \frac{F^{SC}}{Q^{SC}} \quad (27)$$

can be part of an optimal supply commitment. All that matters for incentives is the present discounted value of subsidy payments, not the particular path by which this value is reached.

The limiting case in which an arbitrarily large  $s$  is paid over an arbitrarily short AMC period is of particular interest. The limit is akin to having the donor purchase installed capacity at a linear price  $F^{SC}/Q^{SC}$  per unit. Just as with a perpetual subsidy, which also involves a linear price, the donor behaves like a hybrid between a social planner and a linear-pricing monopsonist, increasingly like the latter the lower its bargaining power. Indeed, the same Lerner index formula from (17) also applies to a supply commitment. To avoid this monopsony distortion and thereby improve efficiency would require some sort of nonlinear contract. We will analyze one sort of

nonlinear contract, a forcing contract, next.

### 3.7. Forcing Contract

A forcing contract pays the subsidy to the firm if and only if it produces a specified level of output. This subsection shows that a forcing contract can be used to improve upon all the policies previously analyzed. Indeed, a suitably designed forcing contract is optimal among the broad class of feasible commitments (feasible in that they satisfy individual rationality for the firm), allowing the donor to attain the integrated outcome.

Consider a forcing contract that pays a subsidy  $s$  per unit from a fund of size  $F$  if and only if the firm's output equals capacity in the integrated benchmark,  $Q^{IN}$ . Supposing for now that the AMC subsidy is temporary, and continuing with the baseline specification that interest accrues in the fund, following (6) the length of the AMC period is given by

$$F = \frac{sQ^{IN}}{r} (1 - e^{-rT}). \quad (28)$$

If the firm rejects the contract, the continuation game resembles that without an AMC analyzed in Section 3.2. Proposition 2 showed the firm's ex ante profit is  $\pi(Q^{NA}, \theta)$  in that case. The firm can earn no more than  $\pi(Q^{NA}, \theta)$  if it accepts the contract but produces a different quantity than  $Q^{IN}$  during the AMC period. Thus it will be an equilibrium for the firm to accept the contract and produce  $Q^{IN}$  if we can show the firm's ex ante profit from the forcing contract is at least  $\pi(Q^{NA}, \theta)$ .

Suppose the firm accepts the forcing contract, builds capacity  $Q^{IN}$ , and produces  $Q^{IN}$  during the AMC period. We will solve for the continuation equilibrium using backward induction, starting from the Nash bargain at  $t = 0$ . The donor's and firm's threat points are 0. Gains from trade equal the present discounted value of joint continuation surplus from the AMC and tail periods:

$$\int_0^T [\alpha B(Q^{IN}) - cQ^{IN} + sQ^{IN}] e^{-rt} dt + \int_T^\infty [\alpha B(Q^{IN}) - cQ^{IN}] e^{-rt} dt = \frac{1}{r} [\alpha B(Q^{IN}) - cQ^{IN}] + F. \quad (29)$$

Implicit in the second term on the left-hand side of (29) is that the firm produces  $Q^{IN}$  in the tail as well as the AMC period. This follows from the principle of full capacity utilization, verified in the proof of the next proposition. The right-hand side of (29) follows from substituting for  $T$  from (28) and rearranging.

Folding the game back, the present discounted value of the firm's ex ante profits equals its share

$\beta$  of (29) less capacity costs. Upon rearranging, this present discounted value is

$$\frac{\beta}{r} \left[ \alpha B(Q^{IN}) - \left( c + \frac{\kappa}{\beta} \right) Q^{IN} \right] + \beta F = \pi(Q^{IN}, \theta) + \beta F. \quad (30)$$

For the forcing contract to work, equation (30) must exceed the firm's ex ante profit  $\pi(Q^{NA}, \theta)$  in the absence of an AMC, which upon rearranging requires  $F \geq [\pi(Q^{NA}, \theta) - \pi(Q^{IN}, \theta)]/\beta$ . The optimal fund size  $F^{FC}$  is just large enough to satisfy this condition with equality. We have the following proposition. The proof in the appendix fills details omitted from the preceding sketch.

**Proposition 7.** *The donor's optimal forcing contract involves a fund of size*

$$F^{FC} = \frac{1}{\beta} \left[ \pi(Q^{NA}, \theta) - \pi(Q^{IN}, \theta) \right], \quad (31)$$

*any subsidy satisfying  $s \geq rF^{FC}/Q^{IN}$ , and a requirement to produce  $Q^{IN}$  during the AMC period. Equilibrium capacity and output under this contract equal  $Q^{IN}$ . This forcing contract is the optimal policy among any that satisfy individual rationality for the firm.*

The optimal subsidy is not uniquely pinned down. An increase in  $s$  reduces the AMC length  $T$ , leaving the present discounted value of the subsidy stream the same at  $F^{FC}$ . The only requirement is that  $s$  is sufficiently high that the firm can extract the whole value of the fund rather than having it continue to grow without bound. If  $s > rF^{FC}/Q^{IN}$ , the subsidy associated with the forcing contract is temporary; if  $s = rF^{FC}/Q^{IN}$ , the subsidy is perpetual; but in either case the present discounted value of subsidy payments is  $F^{FC}$ .

If capacity is contractible, the donor can obtain the outcome from Proposition 7 using a simple capacity-forcing contract, paying the firm  $F^{FC}$  to install capacity  $Q^{IN}$ . Capacity may be difficult to contract on in practice. The firm could sell the factory for another use after collecting the payment or could build a shabby one and abandon it. If capacity is not contractible, the donor would have to resort to the quantity-forcing contract specified in Proposition 7.

### 3.8. Refinements

A variety of additional refinements to the design of a basic AMC were considered in the pilot program. To conserve on space, we have omitted the analysis of these refinements from the text, instead devoting a separate appendix to it. Appendix B examines how the performance of the basic AMC studied so far changes when price caps, purchase guarantees, and country copayments are added. Appendix B also examines the effect of changing the party to which interest accrues

from the AMC fund and of adding a separate player (independent of the donor) who acts as a procurement agent in bargaining, much as GAVI functioned in the pilot program.

## 4. Cost Uncertainty

We have so far studied a model of complete information. Among other things, this meant that the donor was able to design the AMC with symmetric information about the firm's cost parameters  $\kappa$  and  $c$ . In practice, the firm may have private information regarding these parameters, leading to contracting frictions. Our experience on the Economics Expert Group for the pneumococcal pilot bears this concern out: even after considerable expenditure of resources interviewing industry participants, consultants, and our own experts, the Group found it difficult to narrow the range of plausible estimates of  $c$ . With the benefit of time, information, and perhaps some invention, critics have contended that actual costs turned out to be much lower than the operating estimates, charging that the program gave away too much rent to firms (see, e.g., Scudellari 2011).

This section addresses the issue of asymmetric information on AMC design in a simple extension in which production costs can take on two values:  $\underline{c}$  with probability  $\nu \in (0, 1)$  and  $\bar{c}$  with probability  $1 - \nu$ , with  $\underline{c} < \bar{c}$ . Ex ante, the firm knows the realization of  $c$ ; the donor only knows the distribution. Assume that all players learn  $c$  ex post, so that in particular Nash bargaining occurs under symmetric information. Although the pilot program specified a single option, to further understanding here we will examine the design of the optimal screening mechanism, a direct-revelation mechanism offering different forcing contracts depending on the firm's cost announcement. The forcing contract specifies output  $\underline{Q}$  and fund size  $\underline{F}$  if the firm announces cost  $\underline{c}$  and output  $\bar{Q}$  and fund size  $\bar{F}$  if the firm announces  $\bar{c}$ .

This is a fairly standard screening problem, so after introducing some notation, we will jump to the statement of a proposition characterizing the optimum, relegating the details of the analysis to the proof in the appendix. For the required notation, let  $Q^{NA}(c)$  be the equilibrium output in the absence of an AMC, i.e.,  $Q^{NA}(c) = D(c + \kappa/\beta)$ , and  $Q^{IN}(c)$  be the integrated output, i.e.,  $Q^{IN}(c) = D(c + \kappa)$ . Let  $F^{FC}(c)$  be the solution to equation (31) for a particular value of  $c$ . All functions just defined depend on more parameters than  $c$ , but we use the argument to emphasize their dependence on the firm's type. Slightly abusing notation further, we will write the welfare function from (3) as  $w(Q, c)$  and the profit function from (5) as  $\pi(Q, c)$ , the second argument stressing the dependence of both on the parameter  $c$  of interest rather than the whole vector of



model parameters  $\theta$ .

**Proposition 8.** *There exists  $\hat{\beta} \in (0,1)$  such that the following direct revelation mechanism is the optimal AMC in the presence of cost uncertainty.*

- For  $\beta \geq \hat{\beta}$ , each type receives the same forcing contract as under symmetric information: type  $\underline{c}$  receives  $\underline{Q}^* = Q^{IN}(\underline{c})$  and  $\underline{F}^* = F^{FC}(\underline{c})$ , and type  $\bar{c}$  receives  $\bar{Q}^* = Q^{IN}(\bar{c})$  and  $\bar{F}^* = F^{FC}(\bar{c})$ . Expected donor surplus is the same as under symmetric information.
- For  $\beta < \hat{\beta}$ , type  $\underline{c}$ 's quantity remains  $\underline{Q}^* = Q^{IN}(\underline{c})$ , but type  $\bar{c}$ 's is distorted downward to

$$\bar{Q}^* = D \left( \bar{c} + \kappa + \frac{\beta\nu}{1-\nu}(\bar{c} - \underline{c}) \right). \quad (32)$$

Expected donor surplus falls short of that under symmetric information by

$$\nu \left\{ \frac{\beta}{r}(\bar{c} - \underline{c})\bar{Q}^* - \left[ \pi(Q^{NA}(\underline{c}), \underline{c}) - \pi(Q^{NA}(\bar{c}), \bar{c}) \right] \right\} + (1-\nu) \left\{ w(Q^{IN}(\bar{c}), \bar{c}) - w(\bar{Q}^*, \bar{c}) \right\}. \quad (33)$$

The proof in the appendix provides equations for  $\hat{\beta}$  as well as the fund sizes when  $\beta < \hat{\beta}$ . The one wrinkle in the analysis is that the firm's outside option figuring into its participation constraint is type-specific. The firm's outside option is to participate in the no-AMC equilibrium characterized in Section 3.2. The efficient type's profit is higher in this equilibrium than the inefficient type's. For  $\beta$  sufficiently close to 1, the wrinkle indeed matters. The binding constraint for the efficient type is its participation constraint, not its truth-telling constraint. In this case the donor can obtain the integrated outcome in all states by designing the menu options for the two types independently. For  $\beta$  sufficiently close to 0, the wrinkle has no bearing on the problem. We obtain the usual results for screening mechanisms that there is "no distortion at the top" for the efficient type but that the inefficient type's forcing contract is distorted downward to extract information rent from the efficient type. As usual, the binding constraint for the efficient type is its truth-telling constraint.

The analysis points out a political risk associated with the optimal mechanism design. For the case in which  $\beta$  is high, by design, the AMC provides the efficient type with an information rent beyond the already positive profit it would earn in the absence of an AMC. A critic of the AMC program with knowledge of true cost ex post could complain that this rent was an unnecessary giveaway to the firm. The only way to avoid this rent is to distort the inefficient type's quantity down further. This conservative approach to avoiding political risk could inflict substantial social distortions in high-cost states of the world.

## 5. Competing Suppliers

In this section, we show that key results for a monopoly firm are qualitatively similar when the model is extended to allow for competition among  $N \geq 1$  firms. To introduce competition while preserving the feature that ex post prices are set through bargaining, we assume in this section that ex post bargaining is characterized by the Shapley value. When  $N = 1$ , Shapley bargaining reduces to Nash bargaining with  $\beta = 1/2$ . Thus, Shapley bargaining is a generalization of Nash bargaining to multiple parties when these parties have equal bargaining power. Like Nash bargaining, Shapley bargaining presupposes ex post efficiency.

If equilibrium capacity in the absence of an AMC,  $Q_N^{NA}$ , satisfies  $Q_N^{NA} \geq D(c)$ , then there is no hold-up problem and no need for an AMC. Therefore, suppose  $Q_N^{NA} < D(c)$ . Then we can invoke the principle of full capacity utilization. The ex ante profit of firm  $i$ , equal to the present discounted value of its Shapley value each ex post instant less ex ante capacity cost, is

$$\frac{1}{r} \left\{ \sum_{n=1}^N w(n, N) [\alpha B(Q_i + (n-1)Q_j) - \alpha B((n-1)Q_j) - cQ_i] - \kappa Q_i \right\}, \quad (34)$$

Shapley-value coefficient  $w(n, N)$  equals the unconditional probability that a coalition including  $i$ —randomly formed by permuting a grand coalition of  $N$  firms and the donor and taking all agents up to  $i$ —includes exactly  $n-1$  rival firms in addition to the donor.

To derive a formula for  $w(n, N)$ , the numerator of this fraction equals the number of permutations of the players in which  $i$  is preceded by the donor and exactly  $n-1$  rivals, the denominator of which equals the  $(N+1)!$  total permutations of the players. To count the permutations in the numerator, the  $N-1$  rivals can fill the  $N-n$  positions after  $i$  in

$$\frac{(N-1)!}{[(N-1)-(N-n)]!} = \frac{(N-1)!}{(n-1)!} \quad (35)$$

ways. The remaining  $n-1$  rivals and the donor can fill the  $n$  positions preceding  $i$  in  $n!$  ways. Thus,

$$w(n, N) = \frac{[(N-1)!/(n-1)!]n!}{(N+1)!} = \frac{n}{N(N+1)}. \quad (36)$$

Let  $\tilde{w}(n, N)$  be a probability related to  $w(n, N)$  but instead of being unconditional is conditional

on the coalition's including the donor and thus being able to generate a positive surplus. Then,

$$\tilde{w}(n, N) = \frac{w(n, N)}{\sum_{k=1}^N w(k, N)} = 2w(n, N), \quad (37)$$

as

$$\sum_{k=1}^N w(k, N) = \frac{1}{N(N+1)} \sum_{k=1}^N k = \frac{1}{2}. \quad (38)$$

By (37),  $\sum_{n=1}^N \tilde{w}(n, N) = 1$ , so the  $\tilde{w}(n, N)$  are proper weights.

With the formulas for the weights in hand, we can return to the firm's capacity-choice problem. Taking the first-order condition of (34) with respect to  $Q_i$  and imposing  $Q_i = Q_j = Q_N^{NA}/N$  in a symmetric equilibrium yields, after rearranging, the condition for equilibrium industry capacity:

$$\sum_{n=1}^N w(n, N) \left[ P(nQ_N^{NA}/N) - c \right] - \kappa = 0. \quad (39)$$

To express (39) in a form that more readily nests the monopoly case, introduce the concept of the virtual inverse demand under Shapley bargaining:

$$\mathbf{P}(Q, N) \equiv \sum_{n=1}^N [\tilde{w}(n, N) P(nQ/N)]. \quad (40)$$

The reader can verify that  $\mathbf{P}(Q, 1) = P(Q)$ , inverse demand in the monopoly case. Substituting (40) into (39) and rearranging yields

$$\mathbf{P}(Q_N^{NA}, N) = c + 2\kappa. \quad (41)$$

Equation (41) is the equilibrium condition when the principle of full capacity utilization, posited at the outset, holds. The next lemma, proved in the appendix, provides a necessary and sufficient condition for the principle of full capacity utilization to hold, which we can now state concisely in terms of  $\mathbf{P}(Q, N)$ .

**Lemma 2.** *For all  $N \geq 1$ , if*

$$\mathbf{P}(D(c), N) \leq c + 2\kappa, \quad (42)$$

*then the principle of full capacity utilization holds in the absence of an AMC. Condition (42) is non-vacuous, in particular satisfied for sufficiently large  $\kappa > 0$ .*

Condition (42) can be easy to verify in examples. When  $N = 1$ , (42) reduces to  $\kappa \geq 0$ , which is

true in our model regardless of the demand specification. This verifies the principle of full capacity utilization derived in Section 3.3. To take another example, when  $N = 2$  and  $P(Q) = 1 - Q$ , (42) reduces to  $\kappa \geq (1 - c)/12$ .

The proof of the next proposition, provided in the appendix, combines equation (41) with the facts that  $\partial \mathbf{P}(Q, N)/\partial Q < 0$  and  $\mathbf{P}(Q, N) > \mathbf{P}(Q, 1)$  to show that competition leads to more capacity than monopoly in the absence of an AMC.

**Proposition 9.** *For all  $N > 1$ ,  $Q_N^{NA} > Q_1^{NA}$ .*

Proposition 9 says that moving from monopoly to an  $N$ -firm oligopoly increases equilibrium capacity. More generally, one might conjecture that increasing the number of oligopoly firms always increases capacity. This conjecture turns out to be false, as stated in the next proposition. The proof, provided in the appendix, constructs the requisite counterexample.

**Proposition 10.** *There exist cases in which an increase in  $N$  reduces equilibrium capacity in the absence of an AMC.*

The next proposition, proved in the appendix, states that linearity of inverse demand is sufficient for an increase in  $N$  to increase equilibrium capacity starting from any number of firms.

**Proposition 11.** *If inverse demand  $P(Q)$  is linear for all  $Q < D(c)$ , then  $Q_{N''}^{NA} > Q_{N'}^{NA}$  for all natural numbers  $N'' > N'$ .*

Having analyzed the equilibrium with competing firms without an AMC, we proceed to analyze equilibrium when a framework AMC is offered to competing firms. We will focus on the specification that proved useless with a monopoly firm—an AMC involving a temporary subsidy with interest accruing to the escrow—to determine whether competition overturns this result.

Continue to assume that (42) holds, implying that the principle of full capacity utilization holds in the absence of an AMC, so there is scope for an AMC to improve incentives. Posit for now that the condition also holds under the AMC. We will return to verify this below. Given we are analyzing an AMC with a temporary subsidy, the following accounting identity determines the equilibrium duration of the AMC subsidy:

$$F = \frac{s}{r} [Q_i + (n-1)Q_j] (1 - e^{-rT_n}). \quad (43)$$

We have substituted capacities for output because Shapley bargaining is efficient, so firms' total capacity is utilized given the supposition that the principle of full capacity utilization holds. The

subscript on AMC duration  $T_n$  indicates that  $n$  firms ( $i$  and  $n-1$  rivals) participate in the coalition. The Shapley value requires computing  $i$ 's marginal contribution to the coalition's surplus, requiring us to examine the coalition including the  $n-1$  rival firms and the donor but excluding  $i$ . The accounting identity determining AMC duration  $T_{n-1}$  for that coalition is

$$F = \frac{s}{r}(n-1)Q_j(1-e^{-rT_{n-1}}). \quad (44)$$

The ex ante profit of firm  $i$ , equal to the present discounted value of its Shapley value each ex post instant less ex ante capacity cost, is similar to equation (34); but now, in the presence of an AMC, the benefit from the subsidy has to be added to coalitional surpluses:

$$\frac{1}{r} \left\{ \sum_{n=1}^N w(n, N) \left\{ \alpha B(Q_i + (n-1)Q_j) - \alpha B((n-1)Q_j) - cQ_i \right. \right. \\ \left. \left. + s [Q_i + (n-1)Q_j] (1 - e^{-rT_n}) - s(n-1)Q_j (1 - e^{-rT_{n-1}}) \right\} - \kappa Q_i \right\} \quad (45)$$

$$= \frac{1}{r} \left\{ \sum_{n=1}^N w(n, N) [\alpha B(Q_i + (n-1)Q_j) - \alpha B((n-1)Q_j) - cQ_i] - \kappa Q_i \right\} + w(1, N)rF. \quad (46)$$

Equation (46) follows from noting that all positive subsidy terms equal  $rF$  by (43) and (44). Thus, the two subsidy terms in each summand indexed by  $n = 2, \dots, N$  end up canceling each other out. In the summand for the  $n = 1$  index, the second subsidy term is multiplied by  $n-1$ , so equals 0, leaving the first subsidy term, which equals  $rF$ .

Since the last factor in (46) is a constant with respect to  $Q_i$ , the first-order condition from (46) with respect to  $Q_i$  is identical to the first-order condition from (34), implying equilibrium capacity is the same with or without an AMC. Since the principle of full capacity utilization was assumed to hold in the absence of an AMC, the principle carries over to the AMC. This proves the following proposition.

**Proposition 12.** *For all  $N \geq 1$ , if the principle of full capacity utilization holds in the absence of an AMC, equilibrium capacity is the same under a framework AMC with temporary subsidy and interest accruing to the escrow as under no AMC.*

Turn next to the analysis of a supply commitment offered to competing firms. The ex ante profit of firm  $i$  continues to be given by (45), but the AMC duration with a supply commitment is

now given implicitly by the accounting identity

$$F = \frac{s}{r} \min [\hat{q}, Q_i + (n-1)Q_j] (1 - e^{-rT_n}), \quad (47)$$

assuming the principle of full capacity utilization holds. As argued in Section 3.6, the supply commitment is useless unless the target  $\hat{q}$  is set weakly above equilibrium output:

$$Q_i + (N-1)Q_j \leq \hat{q}. \quad (48)$$

Substituting (48) into (47) yields

$$F = \frac{s\hat{q}}{r} (1 - e^{-rT}). \quad (49)$$

The subscript on  $T$  has been dropped recognizing that the lower output associated with coalitions that do not include all the firms do not lengthen the AMC period, which depends on  $\hat{q}$ .

Substituting (49) into (45) and rearranging yields

$$\frac{1}{r} \left\{ \sum_{n=1}^N w(n, N) \left[ \alpha B(Q_i + (n-1)Q_j) - \alpha B((n-1)Q_j) - cQ_i + \frac{rFQ_i}{\hat{q}} \right] - \kappa Q_i \right\}. \quad (50)$$

Taking the first-order condition with respect to  $Q_i$ , imposing symmetry  $Q_i = Q_j = Q_N^{SC}/N$ , imposing that target optimally equals equilibrium capacity  $\hat{q} = Q_N^{SC}$ , and rearranging, yields the equilibrium condition for industry capacity conditional on the AMC terms:

$$\sum_{n=1}^N w(n, N) \left[ P(nQ/N) - c + \frac{rF}{Q} \right] - \kappa = 0. \quad (51)$$

Multiplying through by 2, substituting from (37), and rearranging yields

$$\mathbf{P}(Q_N^{SC}, N) = c + 2\kappa - \frac{rF}{Q_N^{SC}}. \quad (52)$$

This equilibrium condition differs from its analogue, (41), in the absence of an AMC. The proof of the next proposition, provided in the appendix, shows that the difference entails that equilibrium capacity is higher under a supply commitment than no AMC.

**Proposition 13.** *For all  $N \geq 1$ , if the principle of full capacity utilization holds in the absence of an AMC, equilibrium capacity is higher under the optimal supply commitment than under no AMC.*

Thus, the key results we checked carry over from monopoly to competing firms. It continues to be the case that a framework AMC with temporary subsidy and interest accruing to the escrow does not improve incentives but a supply commitment does, implying that a supply commitment is a better design than a framework AMC for any number of firms.

## 6. Design for Technologically Distant Products

In this section we move from an analysis of a technologically close to a technologically distant product. One might naïvely think that the preceding analysis for a close product carries over fairly directly to a distant one with little change beyond the obvious need to scale up the fund  $F$  so that it covers R&D cost  $R$  required for a still-to-be-invented product in addition to the capacity cost  $KQ$ . In fact, the difference between the optimal design of an AMC for close and distant products runs deeper than that. This section will not provide an exhaustive survey of the differences but will focus on some key ones in each of three subsections.

### 6.1. Copayment as “Kill Switch”

This subsection analyzes perhaps the most significant new issue with a technologically distant product. The country copayment, which plays only a supporting role with a close product as discussed further in Appendix B, gains a featured role with distant products. With a close product, there was little danger of wasting AMC funds on a product with limited consumer value because the donor could see the nature of the available products and only commission an AMC if suitable ones happened to exist, writing the product specifications so that only products of sufficient quality were eligible. With a distant product, contractual incompleteness becomes a much larger problem. It may be difficult to write product specifications preventing a low-value product from worming its way into the program. If the donor is allowed an overly strict interpretation of product specifications, it may hold up firms’ investments, undermining the commitment value of the AMC. This is where the country copayment comes in—where the “market” feature of AMCs gains force. Confronting the country with a small copayment creates a kill switch for low-value products which, in the model introduced momentarily, incentivizes the firm to produce high-value products. Giving the kill switch to the donor does not work as well as to the country because the donor cannot commit not to renegotiate with the firm to split the AMC funds locked in the escrow.

Return to the basic model of Section 2; in particular return to the assumption of a monopoly

firm and symmetric cost information (so there is again only one type of production cost,  $c$ ). We will extend the basic model to a distant product as follows. The firm has the opportunity to invest  $R > 0$  in R&D. Doing so results in a useful product providing health benefit  $\alpha B(q)$  and requiring the positive capacity cost  $K$  and production cost  $c$  per unit assumed to this point. If the firm does not invest  $R$ , it produces a useless product providing no health benefit but also involving no associated capacity or production costs. Because of contractual incompleteness, the terms of the AMC written at the distant time before  $R$  needs to be invested cannot distinguish between the useful and useless product.

We seek to analyze the role of the country copayment as a “kill switch” in this setting. For comparison, we first analyze how an AMC without a country copayment performs in this setting. Starting from the design found in Section 3 to be efficient—a forcing contract—we will show this design can be improved on for some parameters by adding a country copayment.

As before, the disposition of unused AMC funds has a crucial impact on incentives. As discussed in the initial setup of the model in Section 2, if the funds flow back to the donor, the AMC has no commitment value because the opportunity cost of the funds will net out in Nash bargaining. Thus, the AMC must specify the funds stay in the escrow. But then the donor gains from bargaining with the firm over trade of even the useless product because they can fulfill the terms of the forcing contract costlessly, dividing the otherwise locked funds  $F$  between them,  $\beta F$  for the firm and  $(1 - \beta)F$  for the donor. To donor cannot commit to pull the “kill switch” on the useless product.

Given the assumption that the firm has no production costs if it does not invest  $R$ , we conclude that the firm’s profit is  $\beta F$  if it does not invest in R&D. If the firm invests, subtracting the additional R&D cost  $R$  from equation (30), we see that it earns  $\pi(Q^{IN}, \theta) + \beta F - R$ . Hence the firm is willing to invest in R&D if and only if  $R \leq \pi(Q^{IN}, \theta)$ . Otherwise, there is no investment, no useful product, and no social surplus. Thus a forcing contract without a country copayment can do nothing to promote R&D incentives.

We will show that efficiency can be improved for a large set of parameters if a small country copayment  $s_c$  is added to the forcing contract. To keep the analysis simple, we will take  $s_c$  to be infinitesimal. Efficiency could be further improved by taking some appropriately calibrated, moderate value of  $s_c$ . The logic of why this is so is discussed in detail in Appendix B: substituting some country for donor funds benefits the donor because it does not internalize the country’s expenditures, just its health benefit. We do not need to complicate the analysis to accommodate



that point here. As long as  $s_c < \alpha_c B(Q^{IN})/Q^{IN}$ , a condition of course satisfied by an infinitesimal copayment, the copayment will not lead the country to demand less than full capacity  $Q^{IN}$ . On the other hand, a positive copayment, no matter how small, leads the country to pull the “kill switch” on the useless product. Now the firm earns 0 if it does not invest compared to the  $\pi(Q^{IN}, \theta) + \beta F - R$  it earns if it does, implying that the firm is now willing to invest if and only if  $R \leq \pi(Q^{IN}, \theta) + \beta F$ . The optimal fund size  $F^{DP}$  (where the superscript indicates the case of a “distant product”) satisfies this condition with equality:  $F^{DP} = [R - \pi(Q^{IN}, \theta)]/\beta$ . The proof (provided in the appendix) of the next proposition shows that the donor benefits from implementing an AMC of this design as long as R&D investment would be efficient for an integrated entity, which is equivalent to the condition  $R < w(Q^{IN}, \theta)$ .

**Proposition 14.** *R&D investment would be efficient for an integrated entity yet cannot be induced by an equilibrium forcing contract without country copayment if and only if  $R \in (\pi(Q^{IN}, \theta), w(Q^{IN}, \theta))$ . If this condition holds, adding an infinitesimal country copayment to the forcing contract not only induces efficient R&D investment but also allow the donor to appropriate all the net surplus from the integrated outcome for itself.*

## 6.2. Extensive-Margin Incentives

As we saw in Section 3, the performance of an AMC for technologically close products depends on the shape of the scheme as much as the level of funding. For example, an AMC involving a forcing contract could induce efficient capacity  $Q^{IN}$  using a fund of size  $F^{FC}$ , while applying those same funds  $F^{FC}$  to a framework AMC with a temporary subsidy would add nothing to incentives for capacity investment (indeed no other funding level would do any better). The nuances of supply commitments and forcing contracts as well as design features such as price caps studied in Appendix B all matter for capacity incentives. With a distant product, the AMC targets R&D as well as capacity investment. In the model, R&D is a fundamentally different type of investment than capacity. Incentives for R&D operate at the extensive margin, depending on the level not the slope of the policy. Incentives for capacity operate at the intensive margin, where slope and other design features become crucial.

We will illustrate these points in this subsection with a stark demonstration. We will take the AMC design found to be useless for a close product—a framework AMC with a temporary subsidy—and show it provides useful incentives for a distant product.

Consider a model close to the one from the previous subsection with a monopoly firm and one cost type  $c$ . Here we will abstract away from the moral hazard problem in the previous subsection,

whereby the firm could save money by producing a useless product. Now assume that there is no choice but to develop a useful product, requiring positive investment  $R$  to do so.

By Proposition 3, a framework AMC with temporary subsidy is useless for capacity incentives. Thus the equilibrium capacity is  $Q^{NA}$  with or without such an AMC. The only possible benefit of such an AMC is to improve R&D incentives. If  $R \leq \pi(Q^{NA}, \theta)$ , then the firm is willing to invest in R&D even in the absence of an AMC. Thus for a framework AMC with temporary subsidy to have the possibility of enhancing R&D incentives relative to no AMC, we must have  $R > \pi(Q^{NA}, \theta)$ . Suppose this inequality holds.

If the firm does not invest in R&D, the firm earns zero profit because nothing is produced. If it invests in R&D, in the presence of a framework AMC with temporary subsidy, its ex ante profit is  $\pi(Q^{NA}, \theta) + \beta F - R$ , which one can see by subtracting the additional R&D cost  $R$  from equation (8). Thus the firm is willing to invest in R&D investment if and only if  $F \geq [R - \pi(Q^{NA}, \theta)]/\beta$ . The efficient fund size satisfies this condition with equality:  $F^{DP} = [R - \pi(Q^{NA}, \theta)]/\beta$ . The proof (in the appendix) of the following proposition shows that the donor strictly benefits from offering a framework AMC with this fund size if and only if  $R < w(Q^{NA}, \theta)$ .

**Proposition 15.** *A framework AMC with temporary subsidy can strictly improve donor surplus relative to no AMC if and only if  $R \in (\pi(Q^{NA}, \theta), w(Q^{NA}, \theta))$ .*

Some remarks about the proposition are in order. The condition on  $R$  in this proposition is identical to that in Propositions 14 except for the capacity used in the right-hand side interval, here  $Q^{NA}$  rather than  $Q^{IN}$ . Proposition 15 does not say the framework AMC is fully efficient. It provides efficient R&D incentives but still does not help with capacity incentives. To improve incentives for capacity investment would require a different AMC design, involving a permanent subsidy, supply commitment, or better yet a forcing contract.

### 6.3. Preferring Distant Products

We saw in the previous subsection that designing an AMC for a technologically distant product can be a conceptually less complex problem than for a close product. In this subsection we show that the problem can be not just conceptually simpler but in fact cheaper for the donor. This is initially surprising because an extra investment  $R$  must be covered for a distant product, an investment assumed already sunk in the case of a close product. The tradeoff is that the problem of asymmetric cost information may be more severe with a close product; with a distant product,

the horizon may be long enough that the firm and donor may be symmetrically uninformed about what costs will turn out to be.

To explore this idea, we start with the model of cost uncertainty from Section 4. This model can be applied without modification to the case of a close product. No additional investment  $R$  needs to be considered because this is assumed to have been sunk before the start of the game. In the case of a distant product, assume that the firm must invest  $R > 0$  for any product to be produced. As in the previous subsection, if this investment is made, the product is certain to provide a health benefit. Assume that the AMC is offered far enough in advance of production that the donor and firm share the same uncertainty about  $c$ , both knowing its distribution— $\underline{c}$  with probability  $\nu \in (0, 1)$  and  $\bar{c}$  with probability  $1 - \nu$ —but not the realization at that point. We will restrict attention to the AMC design found to be efficient in Section 3.7: forcing contracts. Defining the expectations operator as

$$E\left[\pi(Q^{NA}(c), c)\right] = \nu\pi(Q^{NA}(\underline{c}), \underline{c}) + (1 - \nu)\pi(Q^{NA}(\bar{c}), \bar{c}), \quad (53)$$

we have the following proposition.

**Proposition 16.** *Consider the models of a close and distant product, and suppose the donor offers the forcing contract that is efficient in the given model. Let  $\hat{\beta} \in (0, 1)$  be the threshold level of  $\beta$  from Proposition 8.*

- *If  $\beta > \hat{\beta}$ , the donor is indifferent between the close and distant product if*

$$R - E\left[\pi(Q^{NA}(c), c)\right] \quad (54)$$

*is negative and strictly prefers a close product if this expression is positive.*

- *If  $\beta < \hat{\beta}$ , the donor prefers the close product if expression (54) exceeds (33) and the distant product if the reverse inequality holds.*

The appendix provides a formal proof. To derive some intuition, we will examine the cases based on  $\beta$  separately. First suppose  $\beta > \hat{\beta}$ . In that case, as Proposition 8 states, the existence of asymmetric cost information does not harm the donor with a close product. Because the binding constraint for the low-cost type is its participation not truth-telling constraint, the optimal mechanism involves separate first-best forcing contracts for each cost type. With no cost of asymmetric information to trade off against the potentially higher investment cost that needs to be covered, the donor can certainly do no better with a distant product in this case. It is easy to see that must do strictly worse with a distant product for sufficiently high values of  $R$ .

What is less clear is how a distant product can ever achieve a tie with a close product when  $\beta > \hat{\beta}$ , i.e., when the donor suffers no loss due to asymmetric information with a close product. To see how, note that the donor does not have to cover the full R&D cost  $R$  with a distant product but just (54). The firm's ex ante expected profit if it rejects the AMC and engages in Nash bargaining equals  $E[\pi(Q^{NA}(c), c)]$ , which the donor must concede to the firm to ensure it does not reject the AMC contract even in the absence of R&D costs. For a distant product, it is only the excess of  $R$  above  $E[\pi(Q^{NA}(c), c)]$  that must be subsidized to ensure the firm invests in R&D given it is willing to sign the AMC contract.

Next suppose  $\beta < \hat{\beta}$ . Now, by Proposition 8, the presence of asymmetric cost information reduces the donor's surplus by (33). If this loss exceeds the cost (54) of inducing R&D investment with a distant product, then the donor prefers a distant product; if the reverse is true, the donor prefers the close product.

## 7. Conclusion

This paper provides a theoretical analysis of the effectiveness of AMCs in incentivizing investment in new products (such as a vaccine) for low-income countries. We analyze a number of alternative design features, some of which are found to enhance incentives, some of which detract.

Our key contribution may not be particular results about particular design features but rather the broader theoretical framework, which models the AMC as being layered over existing policies that would have proceeded in its absence. In this framework, the incentive effects of an AMC becomes a subtler question. The infusion of substantial AMC funds may merely replace other funds that the donor diverted to other purposes, muting the incentive effects of the AMC. Indeed, we saw that the AMC may have no effect in extreme cases. Incentive power is not a foregone conclusion but requires careful design.

Additional themes emerging from the analysis include that AMC design for a technologically close product (as with pneumococcal vaccine in the pilot) involves quite different considerations than the design for a technologically distant product that is not far along in research and development. Surprisingly, the design for a technologically close product may be more complex. Asymmetry of cost information entails the leakage of information rents to firms, which can be disciplined by possible supply restrictions at some deadweight loss. Manufacturer competition helps expand supply whether or not an AMC is introduced, but competition does interact with the AMC, pro-

moting its effectiveness in a way that is less sensitive to fine design details.

Turning from broad themes to specific implications for the pneumococcal pilot, the finding in Section 5 that, for any number of participating firms, a framework AMC provides no incentives but a supply commitment does supports the Economic Expert Group’s recommendation to switch to a supply commitment. While the framework design could work well in the context of a technologically distant product (the context for which it was originally proposed), the pneumococcal vaccine that was the target of the pilot AMC was a technologically close product.

Turning to the current pandemic, constraints of space and contractual information preclude analysis of the merits of the panoply of contracts offered. Focusing on the contract offered by the U.S. government to Pfizer under Operation Warp Speed—an important example because the product involved turned out to be the first Covid-19 vaccine to succeed in phase-3 trials—the contract exhibits features that theory suggests enhance incentives. The contract was signed early in Pfizer’s development process, before licensure, when considerable investment remained to be incentivized. It resembles a forcing contract, promising around \$2 billion for the first 100 million doses with a call option for millions more (U.S. Health and Human Services 2020). The theory suggests that forcing contracts provide stronger capacity incentives than some other contractual forms.

The analysis in Section 4 highlighted possible political risks associated an AMC for a technologically close product in the presence of asymmetric information about firms’ costs. The donor and country are better off when the supplier turns out to have low cost since the optimal AMC entails large vaccine supply and surplus. In that case, however, the optimal AMC leaves the firm with an information rent—by design—which could be seized on by critics as an unnecessary giveaway to firms, a criticism leveled against the pilot pneumococcal AMC as discussed in the introduction.

## Appendix A: Proofs

This appendix provides omitted proofs from the text of lemmas and propositions.

### Proof of Lemma 1

To apply the inverse function theorem to  $P(q)$ , we need to verify that it is continuously differentiable with nonzero derivative. The assumption that  $B(q)$  is twice continuously differentiable for all  $q \geq 0$  implies  $P'(q) = \alpha B''(q)$  is continuous, in turn implying that  $P(q)$  is continuously differentiable for all  $q \geq 0$ . The assumption that  $B''(q) < 0$  for all  $q \geq 0$  implies  $P'(q) = \alpha B''(q) < 0$ , in turn implying  $P'(q) \neq 0$  for all  $q \geq 0$ . This establishes that  $P(q)$  satisfies the conditions for the inverse function theorem for all  $q \geq 0$ . The theorem then implies that the inverse  $\tilde{D}(P(q)) = P^{-1}(P(q))$  exists for all  $q \geq 0$  and that

$$\tilde{D}'(P(q)) = \frac{1}{P'(q)}. \quad (\text{A1})$$

To bound the codomain of  $P(q)$ , note  $\lim_{q \rightarrow \infty} P(q) = \alpha \lim_{q \rightarrow \infty} B'(q) = 0$ , where the first equality follows from the definition of  $P(q)$  and the second by assumption on the health benefit. Further,  $\lim_{q \rightarrow 0} P(q) = \bar{p}$  by definition. We showed that  $P(q)$  is continuously differentiable for all  $q > 0$ , implying it is continuous for all  $q > 0$ . Thus for all  $p \in (0, \bar{p})$ , there exists  $q > 0$  such that  $P(q) = p$ . Hence for all  $p \in (0, \bar{p})$ ,  $\tilde{D}(p) = \tilde{D}(P(q)) = q > 0$ , and  $\tilde{D}'(p) = \tilde{D}'(P(q)) = 1/P'(q) < 0$  by equation (A1). *Q.E.D.*

### Proof of Proposition 2

The gains from trade over which the donor and firm equal the present discounted value of joint surplus over the continuation game

$$\int_0^{\infty} [\alpha B(q) - cq] e^{-rt} dt = \frac{1}{r} [\alpha B(q) - cq]. \quad (\text{A2})$$

Because Nash bargaining is efficient, parties settle on the  $q$  maximizing (A2) subject to constraint (2). If constraint (2) binds, then  $q^{NA} = Q^{NA}$ . If not,  $q^{NA}$  satisfies the first-order condition of (A2) with respect to  $q$ :  $\alpha B'(q^{NA}) = c$ , or rearranging  $q^{NA} = D(c)$ . Thus,

$$q^{NA} = \min[Q^{NA}, D(c)]. \quad (\text{A3})$$

Given both parties' threat points are 0, the continuation payoff of the firm and donor, respectively, from Nash bargaining are their shares  $\beta$  and  $1 - \beta$ , respectively, of the gains from trade (A2).

Fold the game back to the firm's ex ante choice of capacity. The firm's equilibrium capacity choice  $Q^{NA}$  maximizes the present discounted value of profits

$$\frac{\beta}{r} [\alpha B(q^{NA}) - cq^{NA}] - KQ^{NA} = \frac{\beta}{r} \left\{ \alpha B(\min[Q^{NA}, D(c)]) - c \min[Q^{NA}, D(c)] - \frac{\kappa}{\beta} Q^{NA} \right\}, \quad (\text{A4})$$

where the equality follows from substituting from (A3) and rearranging. Posit that  $Q^{NA} < D(c)$ ;

we will verify this condition later in the proof. Substituting  $Q^{NA}$  for  $\min[Q^{NA}, D(c)]$  in (A4) yields

$$\frac{\beta}{r} \left[ \alpha B(Q^{NA}) - \left( c + \frac{\kappa}{\beta} \right) Q^{NA} \right]. \quad (A5)$$

Differentiating (A5) with respect to  $Q$  yields the first-order condition

$$P(Q^{NA}) - \left( c + \frac{\kappa}{\beta} \right) = 0. \quad (A6)$$

Rearranging,

$$Q^{NA} = D(c + \kappa/\beta). \quad (A7)$$

Above, we posited that  $Q^{NA} < D(c)$ , which by (A7) is equivalent to  $D(c + \kappa/\beta) < D(c)$ . To verify this inequality, there are two cases to consider. If  $c + \kappa/\beta \geq \bar{p}$ , then  $D(c + \kappa/\beta) = 0 < D(c)$ . The last inequality follows from  $0 < c < c + \kappa < \bar{p}$ , implying  $c \in (0, \bar{p})$ , implying  $D(c) > 0$  by Lemma 1. If  $c + \kappa/\beta < \bar{p}$ , then by Lemma 1,  $D'(p) < 0$  for all  $p \in [c, c + \kappa/\beta]$ , implying  $D(c + \kappa/\beta) < D(c)$  by the fundamental theorem of calculus. In either event, we have verified  $Q^{NA} < D(c)$ .

The statement of the proof claims  $Q^{NA} < Q^{IN}$ . By Proposition 1 and equation (A7), this inequality is equivalent to  $D(c + \kappa/\beta) < D(c + \kappa)$ . We omit the verification of this inequality since the argument is similar to the previous paragraph.

Finally, we need to verify the limits. The limit  $\lim_{\beta \rightarrow 1} Q^{NA} = Q^{IN}$  is obvious. To see the other limit,  $\lim_{\beta \rightarrow 0} Q^{NA} = \lim_{p \rightarrow \infty} D(p) = D(\bar{p}) = 0$ . *Q.E.D.*

### Proof of Proposition 3

It remains to show that the principle of full capacity utilization applies to the entire ex post period, comprising the AMC and tail periods. Let  $q^A$  denote equilibrium output during the AMC period. Since Nash bargaining is efficient,  $q^A$  maximizes the ex post joint surplus of the donor and firm, the present discounted value of which is

$$\int_0^T [\alpha B(q^A) - cq^A + sq^A] e^{-rt} dt + \int_T^\infty [\alpha B(Q) - CQ] e^{-rt} dt, \quad (A8)$$

subject to  $q^A \leq Q$ . We argued in the text that the principle of full capacity utilization holds in the tail period, hence the substitution of  $Q$  for output in the second term of (A8). We need to show  $q^A = Q$ .

Before proceeding, we need to provide some analysis of AMC-period length  $T$  appearing in (A8), which is an implicit function of  $q^A$  according to (6). Solving (6) to obtain  $T$  as an explicit function of  $q^A$ ,

$$T = \frac{1}{r} \ln \left( \frac{sq^A}{sq^A - rF} \right). \quad (A9)$$

Differentiating,

$$\frac{\partial T}{\partial q^A} = \frac{-F}{q^A(sq^A - rF)}. \quad (A10)$$

A temporary subsidy satisfies  $sq^A > rF$ , implying (A10) is negative.

Returning to (A8), we will show it is increasing in  $q^A$  for all  $q^A \leq Q$  such that  $Q < D(c)$ , which holds in equilibrium. Differentiating (A8) with respect to  $q^A$ ,

$$\frac{1}{r}[P(q^A) - c + s](1 - e^{-rT}) + sq^A e^{-rT} \cdot \frac{\partial T}{\partial q^A} + \left\{ [\alpha B(q^A) - cq^A] - [\alpha B(Q) - cQ] \right\} e^{-rT} \cdot \frac{\partial T}{\partial q^A}. \quad (\text{A11})$$

The last term is non-negative. To see this, note  $\alpha B(q) - cq$  is concave in  $q$ , maximized for  $q = D(c)$ . This implies  $\alpha B(q^A) - cq^A \leq \alpha B(Q) - cQ$  for  $q^A \leq Q < D(c)$ . Combined with  $\partial T / \partial q^A < 0$ , we have that the last term in (A11) is non-negative. Thus (A11) weakly exceeds

$$\frac{1}{r}[P(q^A) - c + s](1 - e^{-rT}) + sq^A e^{-rT} \cdot \frac{\partial T}{\partial q^A} \quad (\text{A12})$$

$$= \frac{F}{sq^A}[P(q^A) - c], \quad (\text{A13})$$

where (A13) holds by substituting for  $\partial T / \partial q^A$  from (A10); substituting  $s(1 - e^{-rT}) / r = F / q$ , which holds by (6), into the first term in (A12); and substituting  $e^{-rT} = (sq^A - rF) / sq^A$ , which again holds by (6), into the second term in (A12). For  $q^A \leq Q < D(c)$ ,  $P(q^A) > c$ , implying (A13) is positive for all  $q^A \leq Q < D(c)$ . *Q.E.D.*

## Proof of Proposition 5

We first verify the claim in the text that the firm's participation constraint is equivalent to the non-negativity constraint on the subsidy. The firm participates if its ex ante profit under the framework AMC, given in equation (11), weakly exceeds its ex ante profit in the absence of an AMC, given by (4), i.e.,  $\pi(Q, \theta) + \beta sQ / r \geq \pi(Q, \theta)$ , which reduces to  $s \geq 0$ .

We next verify that an implicit solution  $Q^{PS}$  to (16) exists, which is equivalent to the existence of a root  $Q^*$  of the first-order condition (15). We will show that the left-hand side of (15) is positive for  $Q = 0$ , negative in the limit as  $Q \rightarrow \infty$ , so by continuity there exists a root. Since  $B(q)$  is twice continuously differentiable for  $q \geq 0$ ,  $B''(0)$  is finite, implying  $P'(0) = \alpha B''(0)$  is finite. Thus at  $Q = 0$  the left-hand side of (15) is  $P(0) - (c + \kappa) = \bar{p} - (c + \kappa)$ , which is positive by Assumption (1). At the other extreme,

$$\lim_{Q \rightarrow \infty} [P(Q) + \beta Q P'(Q)] \leq \lim_{Q \rightarrow \infty} P(Q) = \alpha \lim_{Q \rightarrow \infty} B'(Q) = 0,$$

where the first step follows from  $P'(Q) < 0$ , the second from the definition of  $P(Q)$ , the third from an assumed property of the health benefit function. Since  $c + \kappa > 0$ , the left-hand side of (15) is negative in the limit as  $Q \rightarrow \infty$ .

With those facts verified, the text preceding the proposition proves that a nontrivial, feasible framework AMC (NFFA) exists if and only if  $s^{PS} > 0$ . We have

$$s^{PS} = c + \frac{\kappa}{\beta} - P(Q^{PS}) = \left( \frac{1 - \beta}{\beta} \right) \kappa + \beta Q^{PS} P'(Q^{PS}), \quad (\text{A14})$$

where the first equality follows from (12) and the second from (16). Taking limits in (A14),



$\lim_{\beta \rightarrow 1} s^{PS} = \lim_{\beta \rightarrow 1} [Q^{PS} P'(Q^{PS})] < 0$ . Thus  $s^{PS} < 0$ —implying a NFFA does not exist—for  $\beta$  sufficiently close to 1. Examining the limit at the opposite extreme,  $\lim_{\beta \rightarrow 0} Q^{PS} = D(c + \kappa) = Q^{IN}$ , implying

$$\lim_{\beta \rightarrow 0} s^{PS} = \kappa \lim_{\beta \rightarrow 0} \left( \frac{1 - \beta}{\beta} \right) + 0 \cdot Q^{IN} P'(Q^{IN}) = \infty.$$

Thus  $s^{PS} > 0$ —implying an NFFA exists—for  $\beta$  sufficiently close to 0.

To establish the non-existence of an NFFA in the limit  $\kappa \rightarrow 0$ , arguments used in the second paragraph of this proof can be used to show that (16) has a finite solution for  $\kappa = 0$ ; call this solution  $Q^0$ . Taking limits in (A14),  $\lim_{\kappa \rightarrow 0} s^{PS} = \beta Q^0 P'(Q^0) < 0$ . Thus  $s^{PS} < 0$ —implying a NFFA does not exist—for  $\kappa$  sufficiently close to 0.

To complete the proof, we establish the remaining claims, which require assumptions on the concavity of the donor's objective function. We will prove the result about the threshold value  $\hat{\beta}$ . The proof for  $\hat{\kappa}$  is similar and omitted. Denote the donor's objective function in (14) by  $\Delta(Q, \theta)$ , where  $\theta$  is the vector of model parameters. The requisite concavity assumption is  $\Delta_{QQ}(Q, \theta) < 0$ , where subscripts indicate partial derivatives. We will show that under this assumption  $\partial s^{PS} / \partial \beta < 0$ . Combined with the limits established above,  $\lim_{\beta \rightarrow 0} s^{PS} > 0$  and  $\lim_{\beta \rightarrow 1} s^{PS} < 0$ , these results are sufficient for the existence of a threshold  $\hat{\beta} \in (0, 1)$  such that  $s^{PS} > 0$  for  $\beta < \hat{\beta}$  and  $s^{PS} \leq 0$  for  $\beta \geq \hat{\beta}$ .

We will use the implicit function rule to derive the sign of  $\partial s^{PS} / \partial \beta$ . The first-order condition  $\Delta_Q(Q^{PS}, \theta) = 0$  gives  $Q^{PS}$  as an implicit function of  $\beta$ . This first-order condition can be transformed so that  $s^{PS}$  appears as an implicit function of  $\beta$ . To do so, solve the incentive-compatibility constraint (12) for  $Q$  as a function of the other variables, obtaining  $Q^*(s, \theta) = D(c + \kappa / \beta - s)$ . For reference, we can sign the following partial derivatives:

$$Q_s^*(s, \theta) = -D' \left( c + \frac{\kappa}{\beta} - s \right) > 0 \quad (\text{A15})$$

$$Q_\beta^*(s, \theta) = -D' \left( c + \frac{\kappa}{\beta} - s \right) \frac{\kappa}{\beta^2} > 0. \quad (\text{A16})$$

Substituting  $Q^*(s, \theta)$  into the donor's first-order condition for  $Q$  gives

$$\Delta_Q(Q^*(s^{PS}, \theta), \theta) = 0 \quad (\text{A17})$$

Applying the implicit function rule to (A17) yields

$$\frac{\partial s^{PS}}{\partial \beta} = - \frac{\Delta_{QQ}(Q^*(s^{PS}, \theta), \theta) Q_\beta^*(s^{PS}, \theta) + \Delta_{Q\beta}(Q^*(s^{PS}, \theta), \theta)}{\Delta_{QQ}(Q^*(s^{PS}, \theta), \theta) Q_s^*(s^{PS}, \theta)}. \quad (\text{A18})$$

All the terms in (A18) have been signed except for one, which we can sign by differentiating (A17) with respect to  $\beta$ , yielding

$$\Delta_{Q\beta}(Q^*(s^{PS}, \theta), \theta) = \frac{1}{r} Q^*(s^{PS}, \theta) P'(Q^*(s^{PS}, \theta)) < 0. \quad (\text{A19})$$

Substituting the signs in (A15), (A16), and (A19) as well as the concavity of  $\Delta(Q, \beta)$  into (A18)

implies  $\partial s^{PS}/\partial \beta < 0$ . *Q.E.D.*

## Proof of Proposition 6

We need to fill in the remaining detail that the fund size  $F^{SC}$  under the optimal supply commitment equals the fund size  $F^{PS}$  under the optimal framework AMC with perpetual subsidy. We first compute  $F^{SC}$ . Substituting equilibrium capacity under the supply commitment, which we showed in the text satisfies  $Q^{SC} = Q^{PS}$ , for  $Q$  in (23) and solving for  $F$  yields

$$F^{SC} = \frac{Q^{PS}}{r} \left[ c + \frac{\kappa}{\beta} - P(Q^{PS}) \right]. \quad (\text{A20})$$

Equations (9) and (12) imply

$$F^{PS} = \frac{s^{PS} Q^{PS}}{r} = \frac{Q^{PS}}{r} \left[ c + \frac{\kappa}{\beta} - P(Q^{PS}) \right]. \quad (\text{A21})$$

Equations (A20) and (A21) together imply  $F^{SC} = F^{PS}$ . *Q.E.D.*

## Proof of Proposition 7

We first verify that the principle of full capacity utilization holds in the tail as well as the AMC period, implying that the firm produces  $Q^{IN}$  throughout the ex post period. The principle holds because the firm's installed capacity satisfies  $Q^{IN} = D(c + \kappa)$  by Proposition 1, and  $D(c + \kappa) < D(c)$ . Efficient Nash bargaining then ensures the firm produces  $q^{IN} = Q^{IN}$ .

To complete the proof, note that by construction, fund size  $F^{FC}$  is sufficient to induce the firm to produce  $Q^{IN}$  during the AMC period under a forcing contract rather than any other output. The firm must build capacity at least  $Q^{IN}$  to produce this much. It is obvious the firm would not build more capacity than  $Q^{IN}$ . Thus the firm's equilibrium capacity and output are  $Q^{IN}$ . This is the same capacity and output as in the integrated benchmark. Therefore the present discounted value of parties' joint surpluses must be the same as in the integrated benchmark, which is the highest possible. By construction of  $F^{FC}$ , the firm's ex ante profit is  $\pi(Q^{NA}, \theta)$ , the least it can earn from any individually rational policy. Thus the forcing contract is optimal for the donor among all policies that are individually rational for the firm. *Q.E.D.*

## Proof of Proposition 8

The optimal direct revelation mechanism is the choice of  $\underline{Q}, \bar{Q}, \underline{F}, \bar{F}$  maximizing expected donor surplus

$$\nu \left\{ \frac{1-\beta}{r} [\alpha B(\underline{Q}) - \underline{c}\underline{Q}] - \beta \underline{F} \right\} + (1-\nu) \left\{ \frac{1-\beta}{r} [\alpha B(\bar{Q}) - \bar{c}\bar{Q}] - \beta \bar{F} \right\} \quad (\text{A22})$$

subject to participation constraints for types  $\underline{c}$  and  $\bar{c}$ , respectively,

$$\pi(\underline{Q}, \underline{c}) + \beta \underline{F} \geq \pi(Q^{NA}(\underline{c}), \underline{c}) \quad (\text{A23})$$

$$\pi(\bar{Q}, \bar{c}) + \beta \bar{F} \geq \pi(Q^{NA}(\bar{c}), \bar{c}) \quad (\text{A24})$$

and truth-telling constraints for types  $\underline{c}$  and  $\bar{c}$ , respectively,

$$\pi(Q, \underline{c}) + \beta \underline{F} \geq \pi(\bar{Q}, \underline{c}) + \beta \bar{F} \quad (\text{A25})$$

$$\pi(\bar{Q}, \bar{c}) + \beta \bar{F} \geq \pi(Q, \bar{c}) + \beta \underline{F}. \quad (\text{A26})$$

The left-hand side of constraints (A23)–(A26) use the expression for ex ante firm profit under a forcing contract provided in (30).

Begin by supposing that constraints (A23) and (A26) do not bind. Then the remaining constraints obviously bind. We can use (A24) and (A25) to solve for  $\beta \underline{F}$  and  $\beta \bar{F}$ . Substituting these values into (A22) leaves the following unconstrained problem in two choice variables,  $\underline{Q}$  and  $\bar{Q}$ :

$$\begin{aligned} & \nu \left\{ \frac{1-\beta}{r} [\alpha B(\underline{Q}) - \underline{c}\underline{Q}] - \pi(Q^{NA}(\bar{c}), \bar{c}) + \pi(\bar{Q}, \bar{c}) - \pi(\bar{Q}, \underline{c}) + \pi(\underline{Q}, \underline{c}) \right\} \\ & + (1-\nu) \left\{ \frac{1-\beta}{r} [\alpha B(\bar{Q}) - \bar{c}\bar{Q}] - \pi(Q^{NA}(\bar{c}), \bar{c}) + \pi(\bar{Q}, \bar{c}) \right\} \end{aligned} \quad (\text{A27})$$

The first-order condition with respect to  $\underline{Q}$  yields  $\underline{Q}^* = D(\underline{c} + \kappa) = Q^{IN}(\underline{c})$ . The first-order condition with respect to  $\bar{Q}$  after rearranging yields the expression in (32). Substituting  $\underline{Q}^*$  and  $\bar{Q}^*$  back into (A24) and (A25) yields

$$\bar{F}^* = \frac{1}{\beta} [\pi(Q^{NA}(\bar{c}), \bar{c}) - \pi(\bar{Q}^*, \bar{c})] \quad (\text{A28})$$

$$\underline{F}^* = \bar{F}^* + \frac{1}{\beta} [\pi(\bar{Q}^*, \underline{c}) - \pi(\underline{Q}^*, \underline{c})]. \quad (\text{A29})$$

We are left to check the ignored constraints. It is easy to see that (A26) is satisfied by the solution. It remains to check (A23). Since (A25) holds with equality and the left-hand sides of (A23) and (A25) are the same, (A23) is satisfied if and only if the right-hand side of (A25) is at least as high as the right-hand side of (A23):

$$\pi(\bar{Q}^*, \underline{c}) + \beta \bar{F}^* \geq \pi(Q^{NA}(\underline{c}), \underline{c}). \quad (\text{A30})$$

Substituting from (A28) for  $\bar{F}^*$  and rearranging, we can write (A30) as

$$0 \leq [\pi(Q^{NA}(\bar{c}), \bar{c}) - \pi(Q^{NA}(\underline{c}), \underline{c})] - [\pi(\bar{Q}^*, \bar{c}) - \pi(\bar{Q}^*, \underline{c})] \quad (\text{A31})$$

$$= \int_{\underline{c}}^{\bar{c}} \frac{d\pi(Q^{NA}(c), c)}{dc} dc - \int_{\underline{c}}^{\bar{c}} \frac{\partial \pi(\bar{Q}^*, c)}{\partial c} dc \quad (\text{A32})$$

$$= - \int_{\underline{c}}^{\bar{c}} Q^{NA}(c) dc + \int_{\underline{c}}^{\bar{c}} \bar{Q}^* dc \quad (\text{A33})$$

$$= \int_{\underline{c}}^{\bar{c}} \left[ D\left(\bar{c} + \kappa + \frac{\nu}{1-\nu} \beta (\bar{c} - \underline{c})\right) - D\left(c + \frac{\kappa}{\beta}\right) \right] dc. \quad (\text{A34})$$

Equation (A32) follows from the fundamental theorem of calculus; (A33) follows from differentiating (5) and applying the envelope theorem; (A34) follows from (32) and the formula  $Q^{NA}(c) =$

$D(c + \kappa/\beta)$ .

Hence (A23) is satisfied if (A34) is non-negative. One can show that (A34) is decreasing in  $\beta$ , positive in the limit as  $\beta \rightarrow 0$ , negative in the limit as  $\beta \rightarrow 1$ . Hence there exists  $\beta \in (0, 1)$  for which (A34) equals 0. This is our threshold  $\hat{\beta}$ .

The proof is completed by deriving the difference in expected donor surplus (33). Expected donor surplus under symmetric information is

$$\nu \left[ w(Q^{IN}(\underline{c}), \underline{c}) - \pi(Q^{NA}(\underline{c}), \underline{c}) \right] + (1 - \nu) \left[ w(Q^{IN}(\bar{c}), \bar{c}) - \pi(Q^{NA}(\bar{c}), \bar{c}) \right] \quad (\text{A35})$$

because for each cost realization  $c$ , the donor can offer a forcing contract allowing it to extract the integrated surplus  $w(Q^{IN}(c), c)$  defined in (3) less the surplus  $\pi(Q^{IN}(c), c)$  that must be offered to the firm to gain its participation. Substituting the equilibrium contract terms into (A22) and subtracting from (A35), after rearranging, yields the difference (33) stated in the proposition. *Q.E.D.*

## Proof of Lemma 2

Let  $\mathbf{D}(P, N)$  be the inverse of  $\mathbf{P}(Q, N)$ , i.e.,  $\mathbf{D}(\mathbf{P}(Q, N), N) = Q$ , which can be interpreted as the virtual demand arising from Shapley bargaining. This virtual demand exists because  $\mathbf{P}(Q, N)$  is a linear combination of invertible functions  $P(Q)$ . Equilibrium quantity in the absence of an AMC can be derived by applying the inverse function to both sides of (41), yielding  $\mathbf{D}(c + 2\kappa)$ . The principle of full capacity utilization is thus

$$\mathbf{D}(c + 2\kappa, N) \leq D(c). \quad (\text{A36})$$

One can apply the inverse function to both sides of (A36) to show it is equivalent to (42).

To show (42) is non-vacuous, by Lemma 1,  $D(c)$  is a finite, positive number under Assumption 1, implying  $\mathbf{P}(D(c), N)$  is as well. Therefore, (A36) holds for sufficiently large  $\kappa$  and is thus non-vacuous. *Q.E.D.*

## Proof of Proposition 9

Differentiating,

$$\frac{\partial \mathbf{P}(Q, N)}{\partial Q} = \frac{2}{N^2(N+1)} \sum_{n=1}^N n^2 P'(nQ/N) < 0,$$

where the inequality holds because  $P'(Q) < 0$  for all  $Q \geq 0$ . Thus

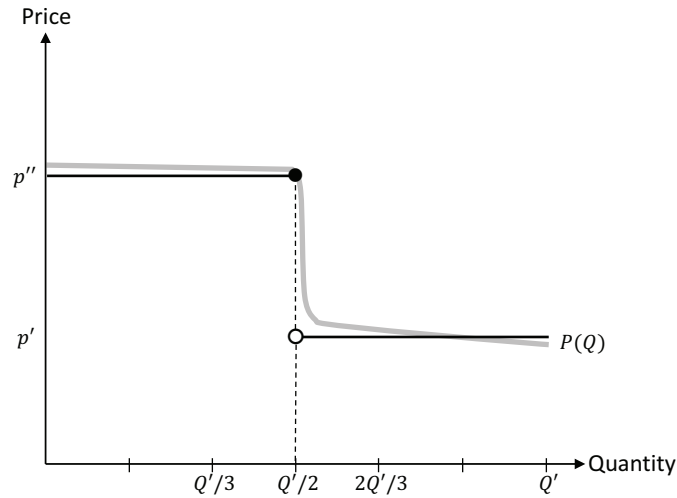
$$\begin{aligned} \mathbf{P}(Q, N) &= \tilde{w}(N, N)P(Q) + \sum_{n=1}^{N-1} \tilde{w}(n, N)P(nQ/N) \\ &> \tilde{w}(N, N)P(Q) + \sum_{n=1}^{N-1} \tilde{w}(n, N)P(Q) \\ &= \sum_{n=1}^N \tilde{w}(n, N)P(Q) \end{aligned}$$

$$\begin{aligned}
&= P(Q) \\
&= \mathbf{P}(Q, 1),
\end{aligned}$$

The second follows from  $P'(Q) < 0$  and  $n/N < 1$  for  $n \leq N-1$ . The remaining lines are either rearrangements or hold by definition. Thus  $\mathbf{P}(Q_N^{NA}, 1) < \mathbf{P}(Q_N^{NA}, N) = c + 2\kappa = \mathbf{P}(Q_1^{NA}, 1)$ , where the inequality was just established and the equalities follow from (41). But  $\mathbf{P}(Q_N^{NA}, 1) < \mathbf{P}(Q_1^{NA}, 1)$  together with  $\partial \mathbf{P} / \partial Q < 0$  implies  $Q_N^{NA} > Q_1^{NA}$ . *Q.E.D.*

### Proof of Proposition 10

The proof proceeds by constructing a counterexample to the conjecture that an increase in  $N$  cannot decrease equilibrium output. The counterexample is based on Figure 3. Consider the inverse demand given by the discontinuous, black curve, equaling  $p''$  for quantities less than or equal to  $Q'/2$  jumping down to  $p'$  for quantities above this threshold, where  $p'' > p'$ . We have  $\mathbf{P}(Q', 2) = (2/3)P(Q') + (1/3)P(Q'/2) = (2/3)p' + (1/3)p''$ , while  $\mathbf{P}(Q', 3) = (1/2)P(Q') + (1/3)P(Q'/2) + (1/6)P(Q'/3) = (5/6)p' + (1/6)p''$ , implying  $\mathbf{P}(Q', 2) > \mathbf{P}(Q', 3)$ . Given this inequality, costs  $c$  and  $\kappa$  can be constructed such that an increase from  $N = 2$  to  $N = 3$  firms reduces equilibrium capacity.



**Figure 3:** Counterexample to competition increasing capacity

While the black inverse demand does not satisfy the properties required of inverse demand such as continuous differentiability, it is clear that an inverse demand such as the grey one in the figure which is close to the black one but does satisfy the required properties can be constructed. *Q.E.D.*

### Proof of Proposition 11

Assuming  $P(Q) = a - bQ$  for  $Q$  in the relevant range,

$$\mathbf{P}(Q, N) = \sum_{n=1}^N \tilde{w}(n, N) \left[ a - b \left( \frac{nQ}{N} \right) \right]$$

$$\begin{aligned}
&= a - \frac{2bQ}{N^2(N+1)} \sum_{n=1}^N n^2 \\
&= a - \left[ \frac{2N+3}{3(N+2)} \right] bQ,
\end{aligned}$$

where the last line follows from  $\sum_{n=1}^N n^2 = n(N+1)(2N+1)/6$ . Similarly,

$$\mathbf{P}(Q, N+1) = a - \left[ \frac{2N+5}{3(N+1)} \right] bQ.$$

Combining these results and rearranging,

$$\mathbf{P}(Q, N+1) - \mathbf{P}(Q, N) = \frac{bQ}{(N+2)(N+1)} > 0.$$

By induction,  $\mathbf{P}(Q, N'') > \mathbf{P}(Q, N')$  for all natural numbers  $N'' > N'$ . Thus  $\mathbf{P}(Q_{N''}^{NA}, N') < \mathbf{P}(Q_{N''}^{NA}, N'') = c + 2\kappa = \mathbf{P}(Q_{N'}^{NA}, N')$ , where the inequality was just established and the equalities follow from (41). But  $\mathbf{P}(Q_{N''}^{NA}, N') < \mathbf{P}(Q_{N'}^{NA}, N')$  together with  $\partial \mathbf{P} / \partial Q < 0$  by the proof of Proposition 9 implies  $Q_{N''}^{NA} > Q_{N'}^{NA}$ . *Q.E.D.*

### Proof of Proposition 13

Combining equations (41) and (52),  $\mathbf{P}(Q_N^{SC}, N) = c + 2\kappa - rF/Q_N^{SC} < c + 2\kappa = \mathbf{P}(Q_N^{NA}, N)$ . We showed  $\partial \mathbf{P}(Q, N) / \partial Q < 0$  in the proof of Proposition 9. Hence,  $Q_N^{SC} > Q_N^{NA}$ . *Q.E.D.*

### Proof of Proposition 14

It remains to show that an efficient forcing contract with infinitesimal country copayment  $s_c$  can deliver the integrated surplus to the donor. We saw from the text that AMC design induced the firm to invest in R&D. The firm's ex ante profit is  $\pi(Q^{IN}, \theta) + \beta F^{DP} - R = 0$  by construction of  $F^{DP}$ . The AMC results in efficient investment and output, generating joint surplus  $w(Q^{IN}, \theta) - R$ , appropriated by the donor. *Q.E.D.*

### Proof of Proposition 15

Assume  $R > \pi(Q^{NA}, \theta)$ . Suppose the donor implements an efficient framework AMC with temporary subsidy, which the text showed involves fund size  $F^{DP} = [R - \pi(Q^{NA}, \theta)] / \beta$ . It remains to derive the condition under which the donor benefits from this policy relative to no AMC. The donor's surplus from this AMC design is

$$\frac{1-\beta}{r} \left[ \alpha B(Q^{NA}) - cQ^{NA} + sQ^{NA}(1 - e^{-rT}) \right] - F^{DP} \quad (\text{A37})$$

$$= \frac{1-\beta}{r} \left[ \alpha B(Q^{NA}) - cQ^{NA} \right] - \beta F^{DP} \quad (\text{A38})$$

$$= \frac{1-\beta}{r} \left[ \alpha B(Q^{NA}) - cQ^{NA} \right] - R + \pi(Q^{NA}, \theta) \quad (\text{A39})$$

$$=w(Q^{NA}, \theta) - R. \quad (\text{A40})$$

Equation (A37) is derived by subtracting the donor's initial expenditure  $F^{DP}$  on the AMC fund from its share  $1 - \beta$  of gains from Nash bargaining from equation (7). We have further substituted  $Q^{NA}$  for  $Q$  following arguments in the text that this design of AMC does not enhance capacity incentives beyond those without an AMC. Equation (A38) follows by substituting from (6), (A39) by substituting for  $F^{DP}$ , and (A40) by first substituting for  $\pi(Q^{NA}, \theta)$  from (5) and then using the definition of  $w(Q, \theta)$  from (3). The donor strictly benefits from this design of AMC if and only if (A40) is positive. *Q.E.D.*

## Proof of Proposition 16

The first step is to derive the optimal forcing contracts for close and distant products. Proposition 8 characterizes the optimal contract for a close product. It remains to derive the optimal contract for a distant product. As with a close product, we will consider a direct revelation mechanism in which the firm announces its cost type before investing in capacity or producing. Quantity and fund size are  $\underline{Q}$  and  $\underline{F}$  if it announces type  $\underline{c}$  and  $\bar{Q}$  and  $\bar{F}$  if it announces  $\bar{c}$ . These terms are set to maximize donor surplus (A22) subject to truth-telling constraints (A25) and (A26) as well as the individual-rationality constraint

$$\nu [\pi(\underline{Q}, \underline{c}) + \beta \underline{F}] + (1 - \nu) [\pi(\bar{Q}, \bar{c}) + \beta \bar{F}] \geq \max \left\{ R, \mathbf{E} \left[ \pi(Q^{NA}(c), c) \right] \right\}. \quad (\text{A41})$$

Following Section 4, we slightly abuse notation by writing the profit function from (5) as  $\pi(Q, c)$ , the second argument stressing its dependence on the parameter  $c$  of interest rather than the whole vector of model parameters  $\theta$ . The two terms in the max operator in (A41) ensure the firm makes two participation decisions correctly, the first ensuring the firm invests in R&D and the second ensuring it accepts the AMC rather than proceeding without one. Because the contract is signed at a point of symmetric uncertainty, and the R&D investment is made before uncertainty is resolved, the individual-rationality constraint need only hold in expectation across states.

We will analyze the following candidate solution (two stars distinguishing the optimum for a distant product from that for a close product). Set outputs at their levels in the integrated optimum:  $\underline{Q}^{**} = Q^{IN}(\underline{c})$  and  $\bar{Q}^{**} = Q^{IN}(\bar{c})$ . Set fund sizes so that they solve the system of equations treating constraints (A25) and (A41) as equalities:

$$\bar{F}^{**} = \frac{1}{\beta} \left\{ \max \left\{ R, \mathbf{E} \left[ \pi(Q^{NA}(c), c) \right] \right\} - \nu \pi(Q^{IN}(\bar{c}), \underline{c}) - (1 - \nu) \pi(Q^{IN}(\bar{c}), \bar{c}) \right\} \quad (\text{A42})$$

$$\underline{F}^{**} = \bar{F}^{**} - \frac{1}{\beta} [\pi(Q^{IN}(\underline{c}), \underline{c}) - \pi(Q^{IN}(\bar{c}), \underline{c})]. \quad (\text{A43})$$

One can verify that these fund sizes satisfy the remaining constraint (A26):

$$\pi(Q^{IN}(\bar{c}), \bar{c}) + \beta \bar{F}^{**} = \pi(Q^{IN}(\bar{c}), \bar{c}) - \pi(Q^{IN}(\bar{c}), \underline{c}) + \pi(Q^{IN}(\underline{c}), \underline{c}) + \beta \underline{F}^{**} \quad (\text{A44})$$

$$= -\frac{\beta}{r} (\bar{c} - \underline{c}) Q^{IN}(\bar{c}) + \pi(Q^{IN}(\underline{c}), \underline{c}) + \beta \underline{F}^{**} \quad (\text{A45})$$

$$= \frac{\beta}{r}(\bar{c} - \underline{c}) [Q^{IN}(\underline{c}) - Q^{IN}(\bar{c})] + \pi(Q^{IN}(\underline{c}), \bar{c}) + \beta F^{**} \quad (\text{A46})$$

$$> \pi(Q^{IN}(\underline{c}), \bar{c}) + \beta F^{**}. \quad (\text{A47})$$

Equation (A44) follows from substituting for  $\bar{F}^{**}$  from (A42), (A45) and (A46) use the definition of the profit function (5), and (A47) follows from  $\bar{c} > \underline{c}$  and  $Q^{IN}(c)$  decreasing in  $c$ . Equations (A44)–(A47) verify (A26) holds.

The proposed solution must solve our proposed constrained optimization problem because it involves outputs maximizing the integrated surplus and forces the individual-rationality constraint to hold with equality. Indeed, the proposed solution must be the optimum over the entire set of contracts—forcing or not, direct revelation or not—satisfying the firm’s individual rationality from a position of symmetric uncertainty. The subsequent analysis will not focus on the terms of this optimal contract but just on the donor’s equilibrium surplus, which equals the integrated surplus less the individually rational ex ante profit for the firm:

$$E[w(Q^{IN}(c), c)] - \max \left\{ R, E \left[ \pi(Q^{NA}(c), c) \right] \right\}, \quad (\text{A48})$$

where  $E[w(Q^{IN}(c), c)]$  is defined analogously to (53).

With the optimal contracts for close and distant products in hand, we turn to a comparison of donor surpluses from the two products. First suppose  $\beta > \hat{\beta}$ . With a close product, by Proposition 8, the donor obtains  $w(Q^{IN}(c), c) - \pi(Q^{NA}(c), c)$  in each state  $c$ , for an expected surplus of  $E[w(Q^{IN}(c), c)] - E[\pi(Q^{NA}(c), c)]$ , which equals (A48) if (54) is negative and exceeds (A48) if (54) is positive.

Next suppose  $\beta < \hat{\beta}$ . With a close product, the donor’s expected surplus is as computed in the previous paragraph less the distortion term (33), denoted  $\xi$  for brevity:

$$E[w(Q^{IN}(c), c)] - E[\pi(Q^{NA}(c), c)] - \xi. \quad (\text{A49})$$

The difference between (A48) and (A49) equals

$$\xi - \max \left\{ 0, R - E \left[ \pi(Q^{NA}(c), c) \right] \right\} = \min \left\{ \xi, \xi - \left\{ R - E \left[ \pi(Q^{NA}(c), c) \right] \right\} \right\}. \quad (\text{A50})$$

Because  $\xi > 0$  when  $\beta < \hat{\beta}$  by Proposition 8, (A50) is positive if (33) exceeds (54) and is negative if (54) exceeds (33). *Q.E.D.*

## Appendix B: Refinements for Technologically Close Products

This appendix analyzes a series of refinements to the design of an AMC for a technologically close product. We examine how the performance of the basic AMC studied in Section 3 changes when various features are added such as price caps, purchase guarantees, and country copayments. We also examine the effect of adding an outside party that collects the accrued interest in the AMC fund or acts as the procurement agent in bargaining.



## B.1. Price Cap

In the original proposals for AMCs (Kremer and Glennerster 2004; Kremer, Barder, and Levine 2005) price caps played a prominent role in AMC design. The idea was to use the AMC subsidy for a dual purpose. In addition to stimulating ex ante investment, the extra surplus provided during the AMC period could be exchanged for lower prices during the tail period, avoiding static deadweight loss then. To streamline the analysis, we have so far ignored price caps but return to them now.

Price caps will turn out to be quite beneficial in our model, but for subtly different reasons than in the original proposals. There is no problem of static deadweight loss in our model because trade is characterized by Nash bargaining, which is ex post efficient. Instead, price caps will help solve hold-up problem. In the absence of a price cap, the firm has an incentive to cut back capacity to put the donor at a bargaining disadvantage by increasing its average value for the product, raising the price the firm can extract from the donor. By limiting this price increase, a price cap reduces the ex ante underinvestment distortion.

Indeed, price caps can be so powerful that an AMC with a suitably designed price cap can completely solve the hold-up problem, achieving the integrated outcome. The efficient policy specifies that the donor pays the firm a lump sum  $F^{FC}$  ex ante in exchange for an agreement that the per-unit price over the whole ex post period be no higher than  $c + \kappa$ . It can be shown that the price cap binds, leading to equilibrium capacity and output  $Q^{IN} = q^{IN} = D(c + \kappa)$ . Lump-sum payment  $F^{FC}$  was fine-tuned in (31) to provide enough surplus to make the firm indifferent between accepting the contract and rejecting it and moving to the subgame without an AMC.

A subtle commitment issue arises with price caps. The firm can threaten not to supply at that price to induce the donor to renegotiate, deleting the price-cap provision. The donor would like to issue a counterthreat not to renegotiate, but because trade requires both parties' assent, this counterthreat would merely enforce the firm's initial threat. To circumvent the commitment problem, the requirement that both parties must assent to trade needs to be altered. The contract could specify a supply guarantee, that the firm must agree to supply a certain amount each instant in exchange for  $F^{FC}$ . The efficient level for supply guarantee is the integrated capacity/output  $Q^{IN}$ .

## B.2. Purchase Guarantee

One modification to the framework AMC that was considered by the Economics Expert Group but ultimately not included in the design of the pneumococcal pilot was a purchase guarantee, i.e., a right given to the firm to be able to sell as much as it wants up to some limit,  $\hat{q}$ . This subsection is devoted to an analysis of how this modification would affect investment incentives under a framework AMC.

A purchase guarantee can be captured in the model by a change in Nash bargaining threat points. In the absence of a purchase guarantee, both bargaining parties must assent to trade, leading to threat points of 0 for both. A purchase guarantee allows the firm the option to force trade. If  $s < c$ , this option is worthless because, if bargaining breaks down, the donor makes no payment above and beyond  $s$ , and  $s$  does not even cover  $c$ . If  $s > c$ , this option is valuable and so would be exercised.

To rule out a proliferation of subcases, before turning to an analysis of the case in which  $s > c$ , we will posit a number of additional conditions. Posit, first, that  $Q \leq \hat{q}$ , i.e., output  $q$  is constrained by  $Q$  rather than  $\hat{q}$ . If  $Q > \hat{q}$ , one can show that the AMC does not provide investment incentives

because the purchase guarantee applies to inframarginal units. Posit, second, that  $F$  is sufficiently small that the subsidy period is temporary. We will explore the case of a permanent subsidy below. The length  $T$  of the temporary subsidy period is determined by (6). Invoking the principle of full capacity utilization, the firm's threat point is

$$\int_0^T (s-c)Qe^{-rt} dt = \frac{(s-c)Q}{r}(1-e^{-rT}) \quad (\text{B1})$$

and the donor's is

$$\int_0^T \alpha B(Q)e^{-rt} dt = \frac{\alpha B(Q)}{r}(1-e^{-rT}). \quad (\text{B2})$$

Paradoxically, despite generating a positive threat point for the firm, the purchase guarantee may worsen the firm's overall bargaining position because the donor's positive threat point may be even better. The firm cannot credibly threaten *not* to exercise its option in a subgame perfect equilibrium, however; so what appears to be a valuable option in isolation may be damaging in a strategic situation.

Indeed, it can be shown that the outcome under a framework AMC with purchase guarantee is worse than without an AMC. Subtracting the sum threat-point surpluses from the present discounted value of joint surplus over the ex post continuation game given in (7) after rearranging yields gains from trade

$$\frac{e^{-rT}}{r} [\alpha B(Q) - cQ]. \quad (\text{B3})$$

Adding the firm's  $\beta$  share of these gains from trade to the firm's threat point gives its Nash bargaining surplus. Subtracting capacity costs yields the following expression for ex ante profits,

$$\frac{\beta e^{-rT}}{r} [\alpha B(Q) - cQ] + \frac{(s-c)Q}{r}(1-e^{-rT}) - KQ \quad (\text{B4})$$

$$= \frac{1}{r} \left\{ \beta \left( 1 - \frac{rF}{sQ} \right) [\alpha B(Q) - cQ] + \frac{s-c}{s} rF - \kappa Q \right\}, \quad (\text{B5})$$

where (B5) follows from substituting for  $T$  from (6). The first-order condition with respect to  $Q$ , after substituting  $P(Q) = \alpha B'(Q)$ , is

$$\frac{\beta}{Q} \left( \frac{rF}{sQ} \right) [\alpha B(Q) - cQ] + \beta \left( 1 - \frac{rF}{sQ} \right) [P(Q) - c] - \kappa = 0. \quad (\text{B6})$$

Substituting  $sQ/(sQ - rF)$  by (6), rearranging, and inverting gives the following expression for equilibrium capacity  $Q^{PG}$  under a purchase guarantee:

$$Q^{PG} = D \left( c + \frac{\kappa}{\beta} - \frac{e^{rT} - 1}{Q^{PG}} \left[ \alpha B(Q^{PG}) - \left( c + \frac{\kappa}{\beta} \right) Q^{PG} \right] \right). \quad (\text{B7})$$

The argument of  $D$  in (B7) is strictly greater than  $c + \kappa/\beta$  if and only if the factor in square brackets is negative, which is the case for  $\beta$  sufficiently close to 0. Thus, if  $D(c + \kappa/\beta) > 0$ , then  $D'(p) < 0$  for all  $p \in (c + \kappa/\beta, \bar{p})$ , implying (B7) is strictly less than  $D(c + \kappa/\beta) = Q^{NA}$ . Hence, for any  $F$  and  $s$  satisfying the posited conditions, and for sufficiently low  $\beta$ , we have the surprising

result that incentives can be impaired by such an AMC.

Adding a purchase guarantee can enhance incentives if layered on a framework AMC with a perpetual subsidy. Indeed, it is easy to construct a policy of this form attaining the integrated benchmark. Set the subsidy as  $s = c + \kappa$  and the limit to the guarantee as  $\hat{q} = Q^{IN}$ . Take the fund size to be large enough that the subsidy is perpetual:  $F \geq (c + \kappa)Q^{IN}/r$  (this is an inequality because a larger fund just means more interest flowing back to the donor, so the funds are not wasted). In equilibrium, there is no Nash bargaining because the firm cannot commit not to exercise its option to produce. To construct the equilibrium, we assume the firm, indifferent as to which capacity it installs in  $[0, Q^{IN}]$ , installs the maximum in this range. The equilibrium is robust in that the donor could always break the firm's indifference by setting  $s$  to be slightly greater than  $c + \kappa$ .

The purchase guarantee layered on a perpetual subsidy functions much like the efficient price cap from the previous subsection. The firm is offered a high enough  $s$  that its cost of capacity and production for the efficient output are covered but no more. The firm does not gain a bargaining advantage from restricting output; there is no need for the donor to bargain with the firm because the firm cannot commit not to chase the subsidy by supplying up to capacity. The subsidy effectively becomes the fixed price in the ex post market.

### B.3. Interest Accrual

The baseline specification behind the policies analyzed so far was that interest accrues in the escrow and, like the principal in the escrow, cannot be used for other purposes. Here we explore two alternatives: interest flows back to the donor or interest flows to some third party. We will see that allowing interest to flow back to the donor completely undermines the investment incentives provided by a framework AMC. In contrast, having the interest flow to a third party helps with incentives. The bargaining parties prefer to get the money out of the fund sooner rather than later, when discounting has destroyed much of its value. The way parties extract money from the fund more quickly is to increase quantity traded, which requires the firm to have installed greater capacity, hence the source of increased incentives. The donor can do better than just giving the interest away to the third party; it can capture the value by charging the third party a lump sum ex ante for the right to collect interest on the fund ex post. In the limit, the donor's surplus approaches that from a framework AMC with interest accruing to the escrow involving a perpetual subsidy.

Because interest does not accrue in the escrow in either alternative, there is no way with a finite  $F$  to endow a perpetual subsidy. Thus, both alternatives for interest accrual must be associated with a temporary subsidy. Both alternatives involve a change to the accounting identity determining the length  $T$  of the AMC period, before given by equation (6), now given by

$$F = sqT. \tag{B8}$$

The identity says that the nominal amount paid into the escrow on the left-hand side equals the nominal amount paid out on the right-hand side.

The change in interest accrual also changes the analysis of Nash bargaining. The donor's threat point is now positive rather than zero because vetoing trade leaves all the money in the escrow from which it earns interest. Trade has the opportunity cost of drawing down the fund, reducing the interest earned by the donor.

To provide the requisite analysis, focus first on the alternative in which interest on the AMC

fund flows back to the donor rather than accruing in the escrow. If bargaining breaks down, the firm obtains no continuation surplus. The donor obtains interest  $rF$  each instant for a present discounted value of  $\int_0^\infty rFe^{-rt} dt = F$ . Thus the donor's threat point is  $F$ .

To compute gains from trade, the present discounted value of joint continuation surplus from trade is

$$\int_0^T [\alpha B(Q) - cQ + sQ]e^{-rt} dt + \int_T^\infty [\alpha B(Q) - cQ]e^{-rt} dt + I. \quad (\text{B9})$$

Capacity  $Q$  has been substituted for output on the basis of the principle of full capacity utilization. The term  $I$  is the present discounted value of the interest earned by the donor, given by the accounting identity

$$F = \int_0^T sqe^{-rt} dt + I. \quad (\text{B10})$$

The left-hand side of (B10) is the present discounted value of payments into the fund, in this case simply the lump-sum endowment. The right-hand side is the present discounted value of all payments out of the fund, both for interest and the subsidy.

Solving (B10) for  $I$ , substituting the resulting expression for  $I$  in (B9), and subtracting off the sum  $F$  of parties' threat point surpluses yields gains from trade

$$\frac{1}{r}[\alpha B(Q) - cQ]. \quad (\text{B11})$$

Adding the firm's  $\beta$  share of these gains and subtracting the investment cost gives the same objective function for the firm's capacity investment as seen in (4) in the absence of an AMC. Equilibrium capacity is thus the same here as without an AMC.

To summarize, we have found that when interest from the AMC fund flows back to the donor rather than staying in the escrow, for all AMC terms  $s > 0$  and  $F > 0$ , the framework AMC adds no incentives for a monopoly firm to invest in capacity. While the result is the same for framework AMCs of finite duration whether interest accrues in the fund or flows back to the donor—both designs useless for incentives—the result holds for different reasons. When interest accrues to the fund, the AMC does contribute to ex post gains from trade. The reason the gains do not contribute to investment incentives is that they do not vary with capacity. When interest flows back to the donor, on the other hand, the AMC contributes nothing to gains from trade. The opportunity cost of \$1 extra subsidy to the firm is \$1 of interest that would have flowed back to the donor. Thus an AMC that lets interest flow out of the fund to the donor does not contribute to incentives for quite stark reasons.

Next consider the alternative in which interest flows to a third party. For concreteness, we will call the third party a bank and suppose it is the same party that sets up and holds the escrow. A crucial assumption is that the bank is a non-strategic player. If the bank were strategic and in particular were able to participate in bargaining ex post, the outcome could collapse to the one just analyzed with interest flowing back to the donor. An added virtue of the analysis of this variant is that it will provide a building block for the next subsection, where we extend the model to the case in which the AMC is set up by a donor but then administered by a different procurement agency, which undertakes ex post bargaining.

Nash bargaining when interest accrues to a third party is similar to that when interest accrues to the escrow in Section 3.4. In both cases, parties' threat points are zero, and the donor does

not directly gain from leaving money in the fund. Thus, we can borrow much of the analysis from Section 3.4 here, in particular the analysis up through the derivation of the left-hand side of equation (8). The analysis diverges when we consider the expression to substitute for the AMC period length  $T$  to derive the right-hand side of (8). Using (B8) rather than (6) to substitute for  $T$ , after rearranging, yields the following expression for the firm's ex ante profits:

$$\frac{\beta}{r} \left[ \alpha B(Q) - \left( c + \frac{\kappa}{\beta} \right) Q + sQ(1 - e^{-rF/sQ}) \right] = \pi(Q, \theta) + \frac{\beta sQ}{r} (1 - e^{-rF/sQ}). \quad (\text{B12})$$

The first-order condition with respect to  $Q$  is

$$P(Q) - \left( c + \frac{\kappa}{\beta} \right) + s - s \left( 1 + \frac{rF}{sQ} \right) e^{-rF/sQ} = 0. \quad (\text{B13})$$

Equation (B13) can be readily compared to the first-order conditions derived in earlier settings. The first three terms,  $P(Q) - c - \kappa/\beta$ , constitute the first-order condition from the benchmark with no AMC. The last two terms reflect the incentive effects provided by the AMC subsidy. Of these two terms, the first, simply  $s$ , is the same incentive effect provided by a perpetual subsidy. The last term reflects the fact that with a fixed fund, an increase in production shrinks the AMC period, dampening the incentives provided by this AMC design. Because interest does not accrue, the current subsidy gain is not completely offset by the future subsidy loss, so the incentive effect does not completely wash out as was seen with the right-hand side of equation (8).

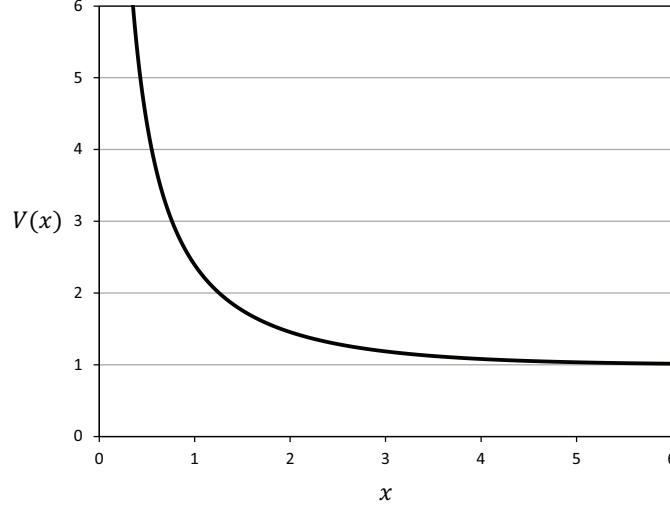
Folding the game back to the donor's design of the optimal AMC, the present discounted value of donor surplus equals its share  $1 - \beta$  of joint continuation surplus (7), less its initial contribution  $F$  to the escrow, plus the bank's lump-sum payment  $I$  from the bank for the right to collect interest ex post:

$$\frac{1 - \beta}{r} [\alpha B(Q) - cQ + sQ(1 - e^{-rT})] - F + I = \frac{1 - \beta}{r} [\alpha B(Q) - cQ] - \frac{\beta}{r} sQ(1 - e^{-rF/sQ}). \quad (\text{B14})$$

The right-hand side follows from substituting from (B10) for  $I$ . This substitution is justified if we assume that banks bid competitively for the right to set up the escrow and collect interest, and these activities are carried out at zero cost. Then  $I$  will equal the present discounted value of payments out of the AMC fund reflected in (B10).

The optimal AMC for the donor is given by the values of  $F$  and  $s$  maximizing (B14) subject to (B13). This constrained optimization problem can be simplified in several steps. First note that  $F$  appears with the same group of terms in (B13) and (B14). We thus substitute the change of variables  $x = rF/sQ$  and maximize with respect to  $x$  rather than  $F$ . Because we can also write  $x = rT$ ,  $x$  can be interpreted as a rescaling of the length of the AMC period. After making the change of variables, we can solve (B13) for  $s$  and substitute out for  $s$  in (B14), yielding

$$\frac{1 - \beta}{r} [\alpha B(Q) - cQ] - \frac{\beta}{r} Q \left[ c + \frac{\kappa}{\beta} - P(Q) \right] V(x), \quad (\text{B15})$$



**Figure 4:** Ratio of marginal cost to marginal benefit of lengthening AMC

where for brevity we have introduced the function  $V(x)$ , graphed in Figure 4, defined as

$$V(x) = \frac{1 - e^{-x}}{1 - (1+x)e^{-x}}. \quad (\text{B16})$$

$V(x)$  can be interpreted as the ratio of marginal cost to marginal benefit (in terms of increased investment incentives) of lengthening the AMC.

We have thus transformed the donor's original constrained optimization problem into the unconstrained problem of choosing  $Q$  and  $x$  to maximize (B15). To complete the solution for the donor's optimum, as the figure shows, and the proof of the next proposition makes rigorous,  $V(x)$  is minimized in the limit as  $x \rightarrow \infty$ , approaching a value of 1. Thus the optimal  $x$  is as large as possible, implying given the change of variables that the optimal AMC lasts as long as possible. Substituting the limit of 1 for  $V(x)$  in (B15), taking the derivative with respect to the remaining choice variable,  $Q$ , and solving for this variable yields the same solution as (16). This establishes the following proposition.

**Proposition 17.** *The donor's surplus from a framework AMC having interest accruing to a third party can come arbitrarily close to that from the optimal framework AMC with interest accrual and a perpetual subsidy, approaching the outcome described in Proposition 4.*

*Proof.* This proof fills in the remaining detail from the analysis leading up to the proposition of verifying the properties of  $V(x)$  seen in Figure 4. In particular, we are left to verify that  $V(x)$  approaches its infimum over the set of  $x > 0$  in the limit as  $x \rightarrow \infty$ . We will do this by showing that  $V'(x) < 0$  for all  $x > 0$  and that  $\lim_{x \rightarrow \infty} V(x) = 1$ . We have

$$V'(x) = \frac{Y(x)e^{-x}}{[1 - (1+x)e^{-x}]^2}. \quad (\text{B17})$$

where  $Y(x) = 1 - x - e^{-x}$ . The sign of  $V'(x)$  is determined by the sign of  $Y(x)$ . We have  $Y'(x) = e^{-x} - 1$ , which is negative for all  $x > 0$ . Given  $Y(0) = 0$ ,  $Y'(x) < 0$  for all  $x > 0$  implies  $Y(x) < 0$  for all  $x > 0$ . Thus  $V'(x) < 0$  for all  $x > 0$ .

To find the limit, rewrite  $V(x)$  as

$$V(x) = \frac{1}{1 - xe^{-x}/(1 - e^{-x})}. \quad (\text{B18})$$

Then

$$\lim_{x \rightarrow \infty} V(x) = \frac{1}{1 - \lim_{x \rightarrow \infty} x/e^x} = \frac{1}{1 - \lim_{x \rightarrow \infty} 1/e^x} = 1, \quad (\text{B19})$$

where the second equality uses l'Hôpital's Rule. *Q.E.D.*

Intuitively, the AMC without interest accrual achieves efficiency by having an increasingly long duration. This helps avoid the perverse incentives associated with AMCs of finite duration, whereby an increase in the firm's capacity and output shortens the AMC period. An increasingly long duration pushes these disincentive effects off into the far future in which they are heavily discounted and thus matter less for ex ante investment. The fund  $F$  needed to endow an arbitrarily long AMC becomes arbitrarily large. The donor is willing to contribute an arbitrarily large  $F$  because the fee  $I$  it earns from the the bank for the right to collect interest ex post also becomes arbitrarily large.

So far we have focused on the effect of interest accrual on the efficiency of framework AMCs. For completeness, it is worth expanding the discussion to other policies including supply commitments and forcing contracts. It is immediate from existing results that any loss from having interest accrue to the donor or third party rather than the escrow can be eliminated in the limit with a suitably designed supply commitment or forcing contract. For both of those policies, we found that the optimal subsidy rate  $s$  was not pinned down. A whole range of  $s$  could equivalently be used to implement the optimum, including an arbitrarily high  $s$ , which in effect results in the whole fund  $F$  being paid out almost in the first instant ex post. Such a policy would render interest payments irrelevant because there would be no time for interest to accrue. Thus it is immediate that any inefficiencies from alternative interest accrual assumptions vanish in the limit for supply commitments and forcing contracts.

## B.4. Procurement Agent

In this subsection, we modify the model to represent the realities of the pneumococcal-vaccine pilot more closely. The donors, the Gates Foundation and five countries, who contributed to the design and funding of the AMC, then stepped aside, passing off the procurement to GAVI. To this point in the analysis, we have modeled those functions as being combined in a player called the donor, but in this subsection we will model them as separate players. The analysis in this section will provide a clearer theoretical understanding of the incentive effects of the pilot's design and whether, for future AMCs, introducing a procurement agent is a design feature that should be copied or avoided.

As before, a donor participates in funding and designing the AMC. It receives flow utility  $\alpha_d B(q) - X_d$  each instant, where  $\alpha_d \in (0, 1)$  represents the degree of donor altruism and  $X_d$  is its expenditure on vaccines. Introduce a new party, a buyer, who takes on the role of the procurement agent, engaging in Nash bargaining with the firm ex post. Its flow utility is  $\alpha_b B(q) - X_b$ , where  $\alpha_b \in (0, 1)$  represents the degree of buyer altruism and  $X_b$  is its expenditure on vaccines. Allowing  $\alpha_d$  to differ from  $\alpha_b$  allows the parties' altruism to differ. Redefine  $\alpha$  as the sum  $\alpha = \alpha_d + \alpha_b$ , and

let  $P(Q) = \alpha B'(Q)$ , i.e.,  $P(Q)$  is inverse demand for the two altruistic parties together. Suppose, as was the case in the pilot, that the buyer participates in the design of the AMC along with the donor. Thus the AMC will be set to maximize their ex ante joint surplus. As was the case in the pilot, suppose interest from the AMC fund accrues to the donor. We will focus on a framework AMC design.

As before, we solve for equilibrium using backward induction. Nash bargaining between the firm and buyer is isomorphic to Nash bargaining between the firm and donor in the variant from the previous subsection in which interest accrued to a third party. In both cases, interest accrues to a party external to bargaining, and so interest does not factor into the bargaining process. Thus, we can borrow much of the analysis from the previous subsection. Because interest does not accrue to the fund, a finite  $F$  cannot fund a perpetual subsidy. AMC duration  $T$  must therefore be finite. The relevant accounting identity determining  $T$  is again given by equation (B8). One can derive parties' Nash bargaining surpluses then fold the game back to the ex ante period, subtracting off the firm's investment cost, to obtain the following expression for the firm's ex ante profits:

$$\frac{\beta}{r} \left[ \alpha_b B(Q) - \left( c + \frac{\kappa}{\beta} \right) Q + sQ(1 - e^{-rF/sQ}) \right]. \quad (\text{B20})$$

This is identical to (B12) except  $\alpha_b$ , the altruism parameter of the buyer who is here the bargaining party, appears in place of  $\alpha$ . The first-order condition with respect to  $Q$  is

$$\alpha_b B'(Q) - c - \frac{\kappa}{\beta} + s - s \left( 1 + \frac{rF}{sQ} \right) e^{-rF/sQ} = 0. \quad (\text{B21})$$

Folding the game back to the design of the optimal framework AMC, in which both donor and buyer participate. The ex ante present discounted value of the buyer's surplus is its  $1 - \beta$  share of the Nash bargaining surplus

$$\frac{1 - \beta}{r} \left[ \alpha_b B(Q) - cQ + sQ(1 - e^{-rF/sQ}) \right]. \quad (\text{B22})$$

The ex ante present discounted value of the donor's surplus is

$$\frac{\alpha_d B(Q)}{r} - F + I = \frac{1}{r} \left[ \alpha_d B(Q) - sQ(1 - e^{-rF/sQ}) \right]. \quad (\text{B23})$$

As an outsider to the Nash bargain, the donor receives flow surplus  $\alpha_d B(Q)$  each instant in effect as an externality. The donor endows the fund with  $F$  but receives the present discounted value of accrued interest  $I$ , which can be derived using accounting identity (B10). Adding (B22) and (B23), substituting for  $I$  from (B10), and rearranging yields the following expression for the ex ante present discounted value of the designers' combined surplus:

$$\frac{1}{r} \left\{ [(1 - \beta)\alpha_b + \alpha_d] B(Q) - (1 - \beta)cQ - \beta sQ(1 - e^{-rF/sQ}) \right\}. \quad (\text{B24})$$

We can solve for the optimal AMC exactly as in the previous subsection, introducing change of variables  $x = rF/sQ$  into (B21) and (B24), solving (B21) for  $s$ , substituting for  $s$  in (B24), factoring out  $V(x)$ , taking the limit  $x \rightarrow \infty$ , which is again optimal here, leading to the limit  $V(x) \rightarrow 1$ . After



those manipulations, we can take the first-order condition of the resulting expression with respect to  $Q$ , providing the following condition for the equilibrium capacity in an optimum:

$$P(Q^{DB}) + \beta\alpha_b Q^{DB} B''(Q^{DB}) - c - \kappa = 0, \quad (\text{B25})$$

where the superscript  $DB$  designating this case refers to the fact that the donor and buyer are separate parties. Equation (B25) can be expressed in a Lerner-index form facilitating comparison with previous results:

$$L^{DB} = \frac{P(Q^{DB}) - (c + \kappa)}{P(Q^{DB})} = \left(\frac{\alpha_b}{\alpha}\right) \frac{\beta}{|\eta^{DB}|}, \quad (\text{B26})$$

where  $\eta^{DB} = Q^{DB} P'(Q^{DB}) / P(Q^{DB})$  is the elasticity of the altruists' combined vaccine demand at the equilibrium quantity.

This analysis suggests that the agency structure used in the pneumococcal pilot, with a separate donor and buyer, contributes to incentives in several ways. First, it has interest accrue to a party not involved in bargaining. This allows a finite-duration AMC to provide incentives just as having interest accrue to a third party did in the previous subsection. But separating donor and buyer goes much further, allowing for stronger incentives than a perpetual AMC did when donor and buyer functions were combined in the donor. Compare the Lerner index formula from that case, (17), to that in (B26). The rightmost-hand of both involve the factor  $\beta/|\eta|$ , but (B26) is multiplied by the fraction  $\alpha_b/\alpha < 1$ . Hence, (B26) is closer to the integrated optimum, 0, than (17). For example, suppose the altruistic parties are symmetric, i.e.,  $\alpha_d = \alpha_b$ , implying  $\alpha_d/\alpha = 1/2$ . Then the Lerner index in (B26) would be halfway between (17) and the integrated optimum.

The intuition for this improvement in incentives is that the hold-up problem turns out to be less severe when one altruistic party carries on the bargaining for many involved in the design of the AMC. There are different lenses through which to view the hold-up problem, but one way to view it is that by underinvesting in capacity, the firm raises its counterparty's value per unit, thus increasing the surplus per unit it can extract from the counterparty. Having multiple counterparties would only multiply this underinvestment distortion. If additional altruistic parties participate ex ante instead of ex post, they can contribute to the committed subsidy, multiplying incentives, without multiplying the underinvestment distortion.

Using this intuition, one can better understand how the altruistic parties would structure the procurement-agency relationship if this were endogenous. Suppose there were two altruistic parties, either or whom could feasibly carry out ex post negotiations, with altruism parameters  $\alpha_1 < \alpha_2$ . Which would be better to designate as the procurement agent? The formula (B26) indicates that the less altruistic one would be better. Lowering the surplus at stake in ex post negotiations reduces the firm's gains from strategically underinvesting relative to a given AMC subsidy, thus reducing the severity of the hold-up problem. In the limit, having a completely non-altruistic procurement agent would lead to the integrated outcome.

## B.5. Country Copayment

The buyer side of the pilot AMC for pneumococcus was yet more complicated in practice than modeled in the previous section. We modeled the buyer side as being divided into two players, a donor which set up the AMC (in practice, the Gates Foundation and country finance ministries)

and a procurement agent which carried out the ex post negotiation (in practice, GAVI). The pilot also involved a third player, the country receiving the vaccine, which was required to contribute a copayment for every dose purchased. Adding a third player to the model on the buyer side—a fourth player in total considering the firm on the supplier side—multiplies the modeling alternatives and clouds the determination of the most natural alternative. We start with a simple alternative that will allow us to highlight some basic issues. The subsection then moves to a more complicated model raising some additional issues. Section 6 highlights an additional first-order effect of country copayments that arises with technologically distant product rather than the technologically close product studied in this section.

Consider a model in which the donor, with altruism parameter  $\alpha_d$ , sets up the AMC. We will consider a framework AMC with temporary subsidy paid at rate  $s_d$ , where the subscript will allow us to distinguish payments coming from the donor-contributed funds from others. We introduce the country receiving the vaccine as a player having the familiar form of utility function used for other players on the buyer side:  $\alpha_c B(q) - X_c$ , where  $X_c$  is its expenditure on vaccines and  $\alpha_c$  is its weight on its own health benefit, inversely related to its marginal utility of income. Parameter  $\alpha_c$  is free in the model but consistent with the donor making purchases on behalf of a very poor country we think of  $\alpha_c$  as being close to 0 and thus the country's marginal utility of income being quite high. The AMC requires the country to make a copayment, denoted  $s_c$ . The country can also engage in Nash bargaining with the firm ex post. To keep the bargaining game simple we suppose that the donor sets up the AMC, which is designed to maximize the joint surplus of donor and country, but does not bargain ex post, leaving the country as the only active party from the buyer side ex post.

By construction, the model is similar to that from the previous subsection with the country now filling the role of the buyer. Besides the notational difference that the country's altruism parameter is  $\alpha_c$  instead of the buyer's  $\alpha_b$ , the only other difference is that the country makes payment  $s_c$  for each unit in addition to any payments resulting from Nash bargaining. However, the  $s_c$  is just an ex post transfer between the bargaining parties. Unlike the subsidy  $s_d$  coming from the AMC fund, which is sunk ex ante, there is nothing sunk about  $s_c$ . It ends up having no bearing on the outcome of Nash bargaining. If the equilibrium transfer from country to firm ends up being higher than  $s_c$ , the country has to top the payment up; if lower, the firm ends up forgiving some of the required payment. The firm's ex ante profit is identical to equation (B20) after substituting  $s_d$  for  $s$  and  $\alpha_c$  for  $\alpha_b$ . The analysis is then identical to that in the previous subsection, leading to the same Lerner index as (B26), with  $\alpha_c$  substituted for  $\alpha_b$ .

Taking  $\alpha_c$  to be close to 0 reveals a benefit of structuring the AMC so that the country make the ex post payments. The Lerner index in (B26) would then be close to 0 and the equilibrium capacity close to the integrated outcome. This is an extension of our earlier finding that having the less-altruistic agent bargain reduces the severity of the hold-up problem. While odd to think of the country as being less altruistic toward itself than an outside party, it is natural to think of a poor country as having a high marginal utility of income. The firm's incentive to distort ex ante capacity investment is reduced when bargaining against a party with a high marginal utility of income. Overall we conclude that a framework AMC with country copayment can enhance efficiency relative to no AMC and for plausible parameters can be quite efficient.

Further efficiencies could be realized by capping the country copayment and tying it to a supply guarantee from the firm. This allows the country copayment to function exactly like a price cap from Section B.1. We saw there that a well-designed price cap can achieve the integrated outcome.

The same can be true here: an appropriately specified copayment can not just approach in the limit  $\alpha_c \rightarrow 0$  but attain the integrated outcome for all  $\alpha_c > 0$ . Specifying the appropriate copayment is a delicate exercise since the country may be on the knife edge between accepting and not. If the copayment is set too high, the country may decline to participate, but reducing it would lead the firm to reject. Lowering the copayment and increasing the size of the AMC fund to compensate the firm could provide the needed cushion.

The caveat that the whole AMC might fail if the copayment is set too high already serves to highlight possible drawbacks with country copayments without the need for complex additional analysis. An additional drawback arises in the more complex model we turn to next. In this model, the donor participates in ex post bargaining in addition to setting up the AMC. Introducing a country copayment entangles the country in ex post bargaining as well, possibly exacerbating the hold-up problem because by withholding the copayment it can veto the firm's receipt of the AMC subsidy.

To this end, consider an alternative model in which the donor designs the AMC ex ante, specifying a fund  $F$  it contributes to, a subsidy rate  $s_d$  from that fund, and a copayment  $s_c$  from the country, which can be toggled on or off. In the design in which the copayment is toggled on, assume that the country must make the copayment as a precondition for release of the AMC subsidy. The analysis is simplest if the AMC involves a perpetual subsidy, so we will consider a framework AMC with a perpetual subsidy with interest accruing to the escrow. The donor, firm, and country engage in ex post bargaining characterized by Shapley value. The weights that the donor and country place on health benefit  $B(q)$  in their utility functions are  $\alpha_d$  and  $\alpha_c$  respectively.

One complexity that arises with Shapley bargaining is that the principle of full capacity utilization may not hold for out-of-equilibrium coalitions with a subset of demanders. Whether or not the principle holds depends on the relative values of  $\alpha_d$ ,  $\alpha_c$ ,  $c$ , and  $\kappa$ , leading to a profusion of subcases. To reduce the profusion, we will treat the case in which  $\alpha_c$  is close to 0, implying that  $\alpha_d$  is close to the sum  $\alpha = \alpha_d + \alpha_c$ . This will ensure that the principle of full capacity utilization applies to any coalition including the donor and firm. To allow for the possibility that the principle does not apply to the coalition including the country and the firm, we introduce some notation. The present discounted value of that coalition's joint surplus ex post is  $[\alpha_c B(q) - cq]/r$ . Ignoring capacity constraints, the output maximizing that surplus satisfies first-order condition  $\alpha_c B'(q) - c = 0$ , implying  $q = (B')^{-1}(c/\alpha_c)$ . Let  $\tilde{Q}$  denote the capacity maximizing that surplus subject to a capacity constraint, i.e.,  $\tilde{Q} = \min[Q, (B')^{-1}(c/\alpha_c)]$ .

An input into the calculation of players' Shapley values, Table B1 lists player's marginal contributions to the coalition including them and all preceding players in the six possible permutations. A player's Shapley value is the average of its entries down a column. The upper panel lists them in the case in which the country is not charged a copayment and the lower panel when it is. The copayment is absent from entries in the upper panel—for obvious reasons since the copayment is not charged there. Notice that the copayment is also absent from entries in the lower panel. The copayment “washes out” of any entry in the lower panel in which it is paid since it is just a transfer, not a marginal contribution to joint surplus.

The only differences between the two panels occur in the third line for permutation DFC and the fourth line for permutation FDC. In each line, a term  $s_d Q$  appearing in another players' marginal contribution in the upper panel, boxed for easy reference, has been reallocated to the country in the lower panel. Coalitions excluding the country lose the AMC subsidy because the country's copayment is a precondition for  $s_d$  to be paid per unit. This makes the country pivotal for the

**Table B1:** Shapley values when country joins donor-firm bargaining

Permutation	Marginal contribution to surplus		
	Firm, F	Donor, D	Country, C
Panel A. Shapley values when country is not charged a copayment			
CDF	$\alpha B(Q) - (c - s_d)Q$	0	0
DCF	$\alpha B(Q) - (c - s_d)Q$	0	0
DFC	$\alpha_d B(Q) - cQ + \boxed{s_d Q}$	0	$\alpha_c B(Q)$
FDC	0	$\alpha_d B(Q) - cQ + \boxed{s_d Q}$	$\alpha_c B(Q)$
FCD	0	$\alpha B(Q) - \alpha_c B(\tilde{Q}) - (c - s_d)(Q - \tilde{Q})$	$\alpha_c B(\tilde{Q}) - (c - s_d)\tilde{Q}$
CFD	$\alpha_c B(\tilde{Q}) - (c - s_d)\tilde{Q}$	$\alpha B(Q) - \alpha_c B(\tilde{Q}) - (c - s_d)(Q - \tilde{Q})$	0
Panel B. Shapley values when country is charged a copayment			
CDF	$\alpha B(Q) - (c - s_d)Q$	0	0
DCF	$\alpha B(Q) - (c - s_d)Q$	0	0
DFC	$\alpha_d B(Q) - cQ$	0	$\alpha_c B(Q) + \boxed{s_d Q}$
FDC	0	$\alpha_d B(Q) - cQ$	$\alpha_c B(Q) + \boxed{s_d Q}$
FCD	0	$\alpha B(Q) - \alpha_c B(\tilde{Q}) - (c - s_d)(Q - \tilde{Q})$	$\alpha_c B(\tilde{Q}) - (c - s_d)\tilde{Q}$
CFD	$\alpha_c B(\tilde{Q}) - (c - s_d)\tilde{Q}$	$\alpha B(Q) - \alpha_c B(\tilde{Q}) - (c - s_d)(Q - \tilde{Q})$	0

Note: All entries are flow surpluses each instant. Dividing by  $r$  converts them into present discounted values.

realization of the  $s_d Q$  term, shifting it into the country's column.

Further analysis is not needed to determine the effect of the copayment. Suppose the AMC is designed by the donor to maximize its ex ante surplus alone. Then the copayment reduces the efficiency of the AMC. The donor loses surplus to the country in the FDC permutation. The firm's loss of surplus in the DFC permutation dulls its marginal investment incentives. Both effects worsen the outcome for the donor. If the AMC is designed by the donor to maximize the combined surplus of donor and country, the copayment policy may be irrelevant. The donor and country offset any ex post transfers between themselves with ex ante transfers. The loss of investment incentives can be offset by increasing  $s_d$  in the optimal AMC design.

## References

- Acemoglu, Daron and Joshua Lim. (2004) “Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry,” *Quarterly Journal of Economics* 119(3): 1049–1090.
- Athey, Susan, Michael Kremer, Christopher M. Snyder, and Alex Tabarrok. (2020) “In the Race for a Coronavirus Vaccine, We Must Go Big. Really, Really Big,” *New York Times* May 4. [Internet.] Downloaded November 23, 2020 from <https://www.nytimes.com/2020/05/04/opinion/coronavirus-vaccine.html>.
- Berndt, Ernst, Rachel Glennerster, Michael Kremer, Jean Lee, Ruth Levine, Georg Weizsäcker, and Heidi Williams. (2007) “Advance Market Commitments for Vaccines Against Neglected Diseases: Estimating Costs and Effectiveness,” *Health Economics* 16(5): 491–511.
- Berndt, Ernst R. and John A. Hurvitz. (2005) “Vaccine Advance-Purchase Agreements For Low-Income Countries: Practical Issues” *Health Affairs* 24(3): 653–665.
- Cernuschi, Tania, Eliane Furrer, Nina Schwalbe, Andrew Jones, Ernst R. Berndt, and Susan McAdams. (2011) “Advance Market Commitment for Pneumococcal Vaccines: Putting Theory into Practice,” *Bulletin of the World Health Organization* 89: 913–918.
- Chalkidou, Kalipso, Hannah Kettler, Ganesh Ramakrishnan, Rachel Silverman, and Adrian Towse. (2020a) “Leave No One Behind: Using a Benefit-Based Advance Market Commitment to Incentivise Development and Global Supply of COVID-19 Vaccines,” Center for Global Development note, Washington DC.
- Chalkidou, Kalipso, Adrian Towse, Rachel Silverman, Martina Garau, and Ganesh Ramakrishnan. (2020b) “Market-Driven, Value-Based, Advance Commitment (MVAC): Accelerating the Development of a Pathbreaking Universal Drug Regimen to End TB,” *BMJ Global Health* 5 (e002061): 1–6.
- Department for International Development. (2009) “Advance Market Commitments for Low-Carbon Development: An Economic Assessment (Final Report),” London.
- GAVI. (2020) “COVAX Explained,” [Internet.] Downloaded November 4, 2020 from <https://www.gavi.org/vaccineswork/covax-explained>.
- Kremer, Michael. (1998) “Patent Buyouts: A Mechanism for Encouraging Innovation,” *Quarterly Journal of Economics* 113(4): 1137–1167.
- Kremer, Michael. (2000a) “Creating Markets for New Vaccines. Part I: Rationale,” *Innovation Policy and the Economy* 1: 35–72.
- Kremer, Michael. (2000b) “Creating Markets for New Vaccines. Part II: Design Issues,” *Innovation Policy and the Economy* 1: 73–118.
- Kremer, Michael and Rachael Glennerster, *Strong Medicine: Creating Incentives for Pharmaceutical Research on Neglected Diseases*. Princeton, NJ: Princeton University Press, 2004.

- Kremer, Michael, Jonathan Levin, and Christopher M. Snyder. (2020) “Advance Market Commitments: Insights from Theory and Experience,” *American Economic Association Papers and Proceedings* 110: 269–273.
- Kremer, Michael and Heidi Williams. (2010) “Incentivizing Innovation: Adding to the Tool Kit,” *Innovation Policy and the Economy* 10: 1–17.
- Levine, Orin, “GAVI Alliance Investment Case: Accelerating the Introduction of Pneumococcal Vaccine into GAVI-Eligible Countries,” Pneumo-ADIP, October 26, 2006.
- Levine, Ruth, Michael Kremer, and Alice Albright. (2005) *Making Markets for Vaccines: Ideas to Action*. Washington, DC: Center for Global Development.
- Light, Donald. (2005) “Making Practical Markets for Vaccines,” *PLoS Medicine* 2(10): 934–938.
- Ozawa, Sachiko, Andrew Mirelman, Meghan L. Stack, Damian G. Walker, and Orin S. Levine. (2012) “Cost-Effectiveness and Economic Benefits of Vaccines in Low- and Middle-Income Countries: A Systematic Review,” *Vaccine* 31(1): 96–108.
- Ridley, David B., Henry G. Grabowski, and Jeffrey L. Moe. (2006) “Developing Drugs for Developing Countries,” *Health Affairs* 25(2): 313–324.
- Scudellari, Megan. (2011) “Are Advance Market Commitments for Drugs a Real Advance?” *Nature Medicine* 17(2): 139.
- Sinha, Anushua, Orin S. Levine, Maria D. Knoll, Tracy A. Lieu. (2007) “Cost Effectiveness of Pneumococcal Conjugate Vaccination in the Prevention of Childhood Mortality: An International Economic Analysis,” *Lancet* 369(9559): 389–396.
- Sonderholm, Jorn. (2011) “A Theoretical Flaw in the Advance Market Commitment Idea,” *Journal of Medical Ethics* 36: 339–343.
- Snyder, Christopher M., Wills Begor, and Ernst R. Berndt, “Economic Perspectives on the Advance Market Commitment for Pneumococcal Vaccines,” *Health Affairs* 30(8): 1508–1517.
- Snyder, Christopher M., Kendall Hoyt, Dimitrios Gouglas, Thomas Johnston, and James Robinson. (2020) “Designing Pull Funding For A COVID-19 Vaccine,” *Health Affairs* 39(9): 1633–1642.
- U.S. Health and Human Services. (2020) “U.S. Government Engages Pfizer to Produce Millions of Doses of COVID-19 Vaccine.” [Internet.] Downloaded November 23, 2020 from <https://www.hhs.gov/about/news/2020/07/22/us-government-engages-pfizer-produce-millions-doses-covid-19-vaccine.html>.
- von Braun, Joachim, et al. (2012) “Innovative Financing for Agriculture, Food Security, and Nutrition,” International Expert Report to Leading Group on Innovative Financing to Fund Development, Paris.
- Weyl, E. Glen and Jean Tirole. (2012) “Market Power Screens Willingness-to-Pay,” *Quarterly Journal of Economics* 127(4): 1971–2003.

World Bank and GAVI. (2006) “Framework Document: Pilot AMC for Pneumococcal Vaccines.” Document prepared for second Donor Working Group meeting November 9, 2006. [Internet.] Downloaded July 3, 2007 from [http://www.vaccineamc.org/files/Framework\\_Pneumo\\_AMC\\_Pilot.pdf](http://www.vaccineamc.org/files/Framework_Pneumo_AMC_Pilot.pdf).

World Health Organization (2007) “Pneumococcal Conjugate Vaccine for Childhood Immunization—WHO Position Paper,” *Weekly Epidemiological Record* 82(12): 93–104.