Increasing the Supply of Health Products in Underserved Regions

Burak Kazaz
bkazaz@syr.edu
Whitman School of Management
Syracuse University
Syracuse, NY 13244

Scott Webster*
scott.webster@asu.edu
W.P. Carey School of Business
Arizona State University
Tempe, AZ 85287

Prashant Yadav
prashant.yadav@insead.edu
Technology and Operations Management, INSEAD, Boulevard Constance, Fontainebleau, France
Center for Global Development, 2055 L Street Washington DC
Global Health and Social Medicine, Harvard Medical School, Boston, MA

September 28, 2023

* Corresponding author

Increasing the Supply of Health Products in Underserved Regions

We study mechanisms that encourage manufacturers of health products to build production and distribution capacity. This is important for low- and middle-income country (LMIC) markets where ability to pay is lower and demand risks are greater. Development finance institutions and philanthropies are beginning to utilize new instruments to incentivize manufacturers to build production/distribution capacity for LMIC markets. The goal of this paper is to understand the effectiveness of such mechanisms in different settings.

We examine four instruments: (1) subsidy proportional to unit sales (sales subsidy), (2) subsidy proportional to unit capacity (variable-capacity subsidy), (3) subsidy proportional to total capacity investment (total-capacity subsidy), (4) a minimum volume guarantee. We analyze incentivized capacity as a function of social-investor budget for each instrument. We show how our framework can be used to identify a social investor's preferred instrument given relevant parameter estimates, and we provide insight into the type of settings where a particular instrument dominates. A sales subsidy dominates when ability to pay is very low; a total-capacity subsidy dominates when ability to pay is low. Outside of these settings, instrument preference is nuanced, though a sales subsidy is dominated by at least one other instrument. When ability to pay is moderate, a variable-capacity subsidy tends to be preferred under high variable-capacity cost and high budget, a volume guarantee tends to be preferred under low variable-capacity cost and high budget, and a total-capacity subsidy tends to be preferred under low budget.

Keywords: global health; social investor; private investment
Received: September 2021; accepted September 2023 by Bradley Staats after four revisions

1. Introduction

This paper examines mechanisms to encourage manufacturers of global health products to build production capacity in order to treat various diseases. We study the impact of alternative actions by a *social investor* that are intended to incentivize a manufacturer to invest in capacity for distribution to under-served markets. A social investor is defined as an investor who uses market-like instruments to achieve social impact returns from their investments. Examples of social investors include governments of developed nations, development finance institutions (e.g., International Finance Corporation, the US DFC, and the European Development Finance Institutions), and philanthropic foundations, (e.g., Bill and Melinda Gates Foundation, Children's Investment Fund Foundation). Development finance institutions and philanthropic foundations are increasingly using incentivizing instruments, in addition to grants, to meet their social/philanthropic objectives (Kania et al. 2015).

Our focus is on existing health products with known efficacy but for which the manufacturer has not invested in production/distribution capacity to serve those in low- and middle-income country (LMIC) markets (e.g., Zambia, Tanzania, Kenya, Senegal, South Africa). One reason for the lack of investment to serve LMIC is low ability to pay,[1] e.g., the price-volume relationship is too low to recover the manufacturer's costs. In developed-country markets, such as the US, EU and Japan, gross margins on most health products are high. It is often argued that this high profit margin is a sufficient incentive for manufacturers to take on these risks and invest in capacity that is optimal for society/patients (at least most times) in developed-country markets, but the same argument cannot be made for LMIC markets. Thus, there is a need to examine incentives to build production/distribution capacity investments specifically for LMIC markets.

Another factor that can contribute to the lack of investment to serve LMIC markets is higher demand risk. LMIC markets generally exhibit higher uncertainty in demand from the manufacturer's perspective compared to developed-country markets for a number of reasons (Kraiselburd and Yadav 2013): (1) historical data on consumption is limited; (2) the data for disease epidemiology is limited and unreliable; (3) funding for health products for some LMIC markets come primarily from external donors (e.g., Global Fund for HIV/AIDS, TB & Malaria, USAID, country payers) with consequent uncertainty in purchase volume (Natarajan and Swaminathan 2014, Rashkova et al. 2017); (4) there is a lack of transparency in processes used by LMIC governments to select reimbursement lists.

Mitigating some of the additional risks is important for improving access to health products in all markets. Collecting better epidemiological data has the potential for reducing demand uncertainty. Some of these activities have already been initiated with financing from organizations such as Unitaid, World Bank, USAID and the Gates Foundation. Aggregating these risks via pooled procurement is another idea that has been implemented by The Global Fund, the Global Drug Facility and USAID's Global Health Programs (Dubois et al. 2019). However, not all this risk can be mitigated or pooled. Because of the difficulty in estimating the intrinsic demand with limited data, payers and manufacturers may have heterogeneous beliefs about demand. In many cases, global donors have better knowledge of intrinsic demand and funding commitments for the disease by country governments (Kraiselburd and Yadav 2013).

In some cases, global donors have a higher tolerance for risk than manufacturers, especially smaller manufacturers. Organizations such as the Bill and Melinda Gates Foundation, Unitaid, UK Foreign and Commonwealth Office (FCDO), CIFF, and British International Investment (BII) have started utilizing risk sharing instruments to incentivize larger investments in capacity to serve LMICs. One specific example is the Bill and Melinda Gates Foundation. With an over $50B endowment the Gates Foundation can take greater risks than a small- or medium-sized health product manufacturer (Bank 2016a). In addition to grants and product-specific subsidies, organizations such as the Gates Foundation, IFC, and US-DFC can offer concessional loans (i.e., low-interest loans) to incentivize a health product manufacturer to invest in

[1] The cost of the health product is borne by end consumers and/or country governments (e.g., government may make the product available to its population at a reduced price).

capacity. They can also offer purchase-volume guarantees to share some of the risks of the manufacturer (Bank 2016a, 2016b). Such mechanisms shift the risk-return tradeoff curve of the manufacturer by either directly lowering the manufacturer's cost of capital or by decreasing their demand uncertainty risk, and thus, resulting in a higher capacity investment. The result could be lower prices and higher manufacturing capacity for LMIC markets.

Typical risk-sharing approaches analyzed in the operations management literature involve agreements such as quantity-flexibility contracts, revenue-sharing contracts, and buyback contracts. This body of literature examines the value of different risk-sharing contracts to one or both parties in a seller-buyer supply chain (e.g., manufacturer selling to a retailer). The insights from this body of literature do not directly translate to our setting for several reasons. First, we take the perspective of a third party interested in influencing a seller-buyer supply chain. The third party is a social investor that wishes to incentivize investment by a health product manufacturer (seller) to serve LMIC governments/residents (buyers). Second, the social investor's objective is to maximize social welfare. The literature on risk-sharing contracts generally focuses on seller and buyer profit (surplus). For health products, there can be a large benefit externality, which is included in social welfare but not in seller-buyer surplus.

In this paper, we study the impact of four types of instruments that a social investor may pursue to incentivize a health product manufacturer to invest in capacity at a desirable level:

(1) Subsidy proportional to sales volume, also known as (aka) sales subsidy
(2) Subsidy proportional to capacity volume, aka variable-capacity subsidy
(3) Subsidy proportional to total capacity investment, aka total-capacity subsidy (or concessional loan)
(4) Minimum volume guarantee

These four instruments reflect the set of viable alternatives that have been explored by agencies such as the Gates Foundation, Unitaid, CIFF, BII, and FCDO. The first three instruments provide a subsidy to the manufacturer and help address the barrier of investment of low ability to pay in LMIC markets. The last instrument shifts some of the risk of uncertain demand to the social investor and helps address the barrier to investment due to demand risk. We want to understand the relative impact of these instruments under differing cost and information structures. The outcome of this research provides a framework that can be used to guide the choice of instruments to achieve development objectives. This work has immediate applications in practice and policy work and advances the field of supply chain finance in the global health sector.

We present and analyze a model that provides conditions under which each instrument is likely to be preferred. These conditions are intertwined and nuanced as they incorporate characteristics of the market, manufacturing costs, and their interactions to incentivize investment in capacity. Our main findings follow. First, when ability to pay is very low, a sales subsidy is the only viable instrument to incentivize investment in capacity. In such a setting, which can arise in LMIC markets, the manufacturer's participation constraint that stems from the fixed cost of building capacity is binding for all other instruments; an output-unit subsidy is necessary to satisfy the participation constraint. The remaining three points apply under the assumption that ability to pay is high enough for other instruments to be viable.

Second, an input-unit subsidy (e.g., variable-capacity subsidy) does not necessarily dominate an output-unit subsidy (e.g., sales subsidy). If there is no participation constraint (e.g., no fixed cost), then our findings align with a finding in the literature: a variable-capacity subsidy incentivizes greater investment in capacity for a given budget than a sales subsidy; the leverage from input-unit subsidy is higher because the subsidy affects both overage and underage cost whereas an output-unit subsidy only affects underage cost. However, in the presence of a participation constraint, the "extra cost" of a sales subsidy becomes an advantage; it can cover the manufacturer's fixed cost, thereby incentivizing investment that would not occur under a variable-capacity subsidy (e.g., when budget is low). When the budget is high enough that a variable-capacity subsidy becomes viable (i.e., participation constraint is no longer binding), then the variable-capacity subsidy dominates the sales subsidy.

Third, a total-capacity subsidy and a sales subsidy exhibit a structural similarity that manifests in two fundamental properties: (1) the minimum budget that incentivizes investment in capacity (*threshold* budget) is the same for both instruments, (2) at any budget above the threshold, the total-capacity subsidy yields higher investment in capacity than a sales subsidy. The total-capacity subsidy includes a subsidy on input units (greater leverage compared to an output-unit subsidy) as well as a subsidy on fixed cost to address the participation constraint. However, the contribution to fixed cost means that a total-capacity subsidy is dominated by a variable-capacity subsidy when the budget is high enough that the participation constraint under a variable-capacity subsidy is not binding.

Fourth, the viability of a volume guarantee is linked to breakeven volume, i.e., fixed cost of capacity divided by the contribution margin (including variable cost of capacity). The main managerial consequence is that a volume guarantee is likely to be dominated by another instrument unless ability to pay (and consequently, contribution margin) is moderate-to-high.

After summarizing relevant literature in the next section, we present and analyze a model for the social investor and the manufacturer in Section 3. Section 4 describes and analyzes the different instruments for incentivizing investment in capacity. Section 5 presents analysis of illustrates the application of our model to three real-world cases. Section 6 summarizes our main findings, with emphasis on implications for social investors. Derivations and proofs are available in an online appendix.

2. Related Literature

There is a vast literature that examines the impact of market interventions by government or nongovernmental organizations to improve social welfare. Areas of application include agriculture (e.g., Stiglitz 1987, Tang et al. 2015), clean or efficient energy (e.g., Aldy et al. 2019, Alizamir et al. 2016, Cohen et al. 2016, Krass et al. 2013, Raz and Ovchinnikov 2015, Yu et al. 2018), medicines (e.g., Arifoğlu et al. 2012, Chick et al. 2008, Kazaz et al. 2016, Taylor and Xiao 2014, Park et al. 2018, Martin et al. 2020), and remanufacturing (e.g., Atasu and Subramanian 2012, Atasu et al. 2019, Mitra and Webster 2008, Webster and Mitra 2007). There is also a wide literature on the timing and the amount of capacity expansion under demand uncertainty (e.g., Van Mieghem 2003, Erkoc and Wu 2005, Ozer and Wei 2006, Song et al. 2020). The models in this literature often exhibit a newsvendor structure, as is the case for the capacity investment decision model that we present in Section 3.3.

We consider four possible instruments to incentivize investment in product/distribution capacity: sales subsidy, variable-capacity subsidy, total-capacity subsidy, and volume guarantee. We note that economics literature has examined the relative effects of a subsidy on input units (e.g., variable-capacity subsidy) versus a subsidy on output units (e.g., sales subsidy). For example, Parish and McLaren (1982) provide an analytical treatment and show neither subsidy dominates the other, and Aldy et al. (2019) empirically find that an output subsidy is more cost-effective for promoting wind energy. Berndt et al. (2007) and Kremer et al. (2022) examine advance market commitment (AMC) to vaccines in the form of a minimum price to pay per immunized person; this corresponds to a sales subsidy, which is the focus of the AMC literature. This stream of economics literature has not examined the effects of subsidies in a setting with risk stemming from demand uncertainty. There are relatively few papers that have considered one or more of these interventions under demand risk. In the following, we summarize this literature and clarify similarities and differences relative to our work.

Several earlier publications focused on determining whether input or output subsidies, or their combination, are beneficial in different settings. Taylor and Xiao (2014) consider the problem of donor seeking to maximize the expected unit sales of a malaria medicine (per period) subject to a budget constraint. The donor considers two types of retailer subsidies to influence optimal ordering/pricing policies to improve social welfare: an input subsidy (per unit purchased by the retailer) and an output subsidy (per unit sold by the retailer). The authors show that, while the optimal order quantity is more sensitive to the input subsidy and the optimal retail price of the medicine is more sensitive to the output subsidy, it is optimal to only use an input subsidy. The finding relies on the fact that an input subsidy reduces both the cost of understocking (loss from revenues) and the cost of overstocking (loss from too much purchased). Raz and Ovchinnikov (2015) show similar findings in the design of incentives offered by government to maximize social welfare of a product in the context of electric vehicles. They consider two types of incentives: (1)

subsidy paid to the manufacturer for each unit produced (input subsidy paid to the manufacturer), (2) rebate paid to the consumer for each unit purchased (output subsidy paid to the consumer). They show that an output subsidy can co-exist with an input subsidy in the optimal solution that maximizes social welfare. Cohen et al. (2016) extend this work by examining the use of a manufacturer cost subsidy and a consumer rebate to increase the adoption of electric vehicles. They focus on understanding the impact of ignoring demand uncertainty when setting subsidies and show numerically that neither the input nor output subsidy universally dominates the other.

The paper most closely related to our work is Martin et al. (2020) who study optimal incentive instruments for increasing the supply of vaccines in developing countries. They study three types of subsidy instruments: (1) a subsidy on sales up to a sales volume cap, (2) a capacity-dependent subsidy on sales where the per-unit subsidy increases if capacity is above a threshold, (3) a subsidy on sales with an additional subsidy payment for each unit of unused capacity. The first two instruments are variations of an output subsidy, and the third instrument includes both input (payment proportional to capacity) and output (payment proportional to sales) subsidies.

In summary, our study differs from the above publications because we consider two different input subsidies and one risk-shifting instrument as well as an output subsidy. In the setting of our study, the fixed cost of capacity plays a meaningful role and offers a unique feature that has not been considered in the literature. The presence of a fixed cost introduces manufacturer participation constraint. More specifically, we consider a sales subsidy and two types of capacity subsidies—payment proportional to the variable cost of capacity (variable-capacity subsidy) and payment proportional to the total cost of capacity (total-capacity subsidy). To our knowledge, the total-capacity subsidy has not previously been considered and becomes relevant because of the manufacturer participation constraint. In addition, we consider a volume guarantee that shifts demand risk from manufacturer to the social investor and is distinct from classical input/output subsidy instruments. Finally, we remark that the presence of a manufacturer participation constraint leads to nuanced conclusions on instrument preference relative to the literature. For example, the general finding from the literature that an input subsidy is preferred over an output subsidy no longer holds. This is due to the presence of the manufacturer's participation constraint. Furthermore, both the total-capacity and the volume guarantee instruments offer unique advantages over the traditional input and output subsidies in some settings. For example, the total capacity subsidy dominates when ability to pay is low, and the volume guarantee is more likely to be preferred under moderate ability to pay when variable-capacity cost is low.

Each of the instruments we study are considered by practitioners. Our objective is to provide insight that can help guide the social investor on the choice of instrument. We identify conditions under which each instrument is more cost effective than the others. We focus on the implementation of each instrument in isolation for two reasons. First, our interactions with those who have worked on implementing these instruments indicate that the transactional and contractual costs of combined instruments are significantly high. These social investors indicate that, unless the benefits of combined instruments are very high, pure instruments are more suitable in practice. Second, the simultaneous optimization of parameters across multiple instruments adds significant complexity warranting a separate study. Our paper provides an important step to this line of research as the study of instrument combinations relies on a keen understanding of factors that influence the performance of each instrument.

3. Social Investor and Manufacturer Models

The social investor plays the role of a principal seeking to incentivize an agent (manufacturer) to invest in production/distribution capacity to serve the LMIC markets. Section 3.1 describes characteristics of the problem in practice. The content in this section provides background and context for models presented in sections 3.2 and 3.3.

3.1. Examples from Practice and Problem Characteristics

Our study draws on interactions with those who have worked on projects to incentivize investment in capacity to serve LMIC markets, including the Jadelle contraceptive implant, the Hologic viral load test, and next generation long-lasting insecticide-treated bed nets. We begin with a brief description of each of these examples. Then we summarize the common features that form the basis of our model.[2]

Contraceptive implants have been widely available in developed country markets (e.g., Norplant, the first contraceptive implant, became available in 1983). The 2012 price for the Jadelle implant was $18/unit, approximately twice the upper limit on the affordable price in LMIC markets. A coalition of social investors worked with Bayer to incentivize investment in a new production line for the Jadelle implant to serve these markets at a price of $8.50/unit. While subsidy instruments were considered, the group provided a volume guarantee of 27 million units to be supplied to 50 LMICs over a six-year period. A similar agreement was reached later with Merck, the manufacturer of another implant, Implanon.

HIV-AIDS viral load testing is essential for bringing viral load below a threshold at which HIV is no longer transmitted; it is a critical step toward achieving the UN goal of eliminating the HIV-AIDS epidemic by 2030. The availability of test devices/materials has been limited in many African countries due to high cost. CHAI conducted a study of viral load test cost structure and LMIC ability to pay. Through a three-year volume guarantee agreement involving CHAI, MedAccess, and supplier Hologic, viral load

[2] One of the authors has worked for a nonprofit social investor. Our study draws on a sound understanding of real-world challenges and approaches for harnessing the private sector to improve health outcomes through strategic investments.

test instruments have been installed throughout African countries. Viral load tests are provided by Hologic at a price of $12/test.

Long-lasting insecticide-treated bed nets (LLINs) help prevent the spread of malaria. In 2017, the WHO noted growing resistance to pyrethroids, the only WHO-approved insecticide for LLINs, and prioritized the development of the next generation LLIN at a low cost (e.g., at a price comparable to the price of pyrethroid-LLINs). To incentivize investment, WHO and partners including The Global Fund, Gates Foundation, Unitaid, USAID, President's Malaria Initiative considered paying the successful supplier(s) the difference between next generation LLINs and pyrethroid-LLINs as tests/trials take place (ten Brink et al. 2018), i.e., a sales subsidy. Development and trials are ongoing.

The above examples illustrate characteristics of settings that motivate this research. The social investor draws on established relationships with LMIC governments/agencies and other resources to determine (1) that there is a significant and pressing need for a health product relative to current availability and (2) that governments and agencies are open to efforts to increase the supply of the product in their markets. The social investor has conducted research to understand the product's cost structure, the willingness/ability of country governments to pay for the health product (i.e., upper limit on price), and an estimate of annual volumes consistent with the goal of maximizing social welfare. In practice, the social investor recognizes that volume projections are not precise (e.g., benefit externality, the level of need, and other parameters are challenging to estimate). In recognition of this characteristic, our goal is to develop a framework that can help a social investor to gain insight into the cost of instruments to incentivize capacity investment over a range of different volume levels.

3.2. Capacity Investment from the Social Investor's Perspective

The social investor has identified a high priority need for a health product that is insufficiently available in LMIC markets, including an estimate of "ideal" annual volumes. In this section, we describe factors that a social investor considers when developing volume estimates. We then define a model that specifies the micro-level decision process that underlies a social investor's preferred capacity. We use the model to identify an upper limit of incentivized capacity in our numerical illustrations in Section 5.

Social welfare is the difference between the value of the health product to those in need and the cost to provide the health product. Obtaining accurate estimates of health product value, cost structure, and market need can be challenging (Levine et al. 2008). However, recent initiatives led by organizations like UNITAID, CHAI, and USAID have begun systematically forecasting global demand for certain health products, particularly those funded by international agencies. Two sources of cost information are available. First, academic production engineers have developed modeling-based estimation methods for both fixed and variable production costs (Basu et al. 2008, Hill, 2018). Second, as major social investors like the Gates Foundation and MedAccess increasingly use the instruments discussed in this paper, they have started funding specialized organizations like CHAI to estimate production cost curves. The societal value resulting from the use of health products is estimated by global Health Technology Assessment (HTA) agencies for various new health products (Tantivess et al. 2017). Obtaining accurate estimates of need, costs, and health system/societal value is much more challenging for products that are either new and innovative or when financing and purchasing are carried out directly by country governments rather than international agencies.

The above discussion summarizes considerations that underlie the social investor's volume projections that we formalize through a model. We identify socially optimal capacity to serve needs over a finite period of time that we refer to as the *investment horizon* (e.g., several years). Details on the duration and meaning of the investment horizon appear in Section 4; for now, we simply note that it is the relevant duration for evaluating a capital investment under consideration.

The number of individuals that benefit from use of the health product during the period is uncertain. The social investor's demand forecast is μ, which is the expected value of uncertain demand $\tilde{d}$. Let c_v denote the marginal cost of production, c_k denote the marginal cost of capacity, c_f denote the fixed cost of capacity, and υ denote the social value per unit of the health product that includes externality benefit where $\upsilon > c_v + c_k$ (if the inequality did not hold, then the social investor would not be interested in incentivizing investment in capacity). The social welfare from capacity decision x during the period is the difference between value and cost:

$$\Pi(x) = E\left[(\upsilon - c_v)\min\left\{\tilde{d}, x\right\} - c_k x - c_f\right].$$

Note that the social welfare function has a newsvendor structure, which is evident in the following characterizations of socially optimal capacity, denoted x_1^*.

Proposition 1. *Socially optimal capacity* x_1^* *satisfies* $\Pr\left(x_1^* > \tilde{d}\right) = c_k / (\upsilon - c_v)$.

(The proof of Proposition 1 is straightforward and is omitted.) In any realistic setting, $\Pi\left(x_1^*\right) > 0$; otherwise, the social investor is not interested in incentivizing the manufacturer to invest in capacity.

3.3. Capacity Investment from the Manufacturer's Perspective

The manufacturing firm does not currently supply the health product in LMIC markets. The focus of our model is the level of capacity the manufacturer will build as a function of incentives provided by the social investor. We present a base model of the manufacturer's decision process in this section, then augment this model to accommodate alternative incentive instruments in Section 4.

For capital investment decisions, firms evaluate the payoff from the investment's financial flows over a fixed period. As noted in the previous section, we refer to this length of time as the investment horizon.

After building capacity (or retrofitting an existing facility), the manufacturer produces according to observed demand over the investment horizon up to available capacity. Figure 1 presents the timeline of decisions and events in our model.

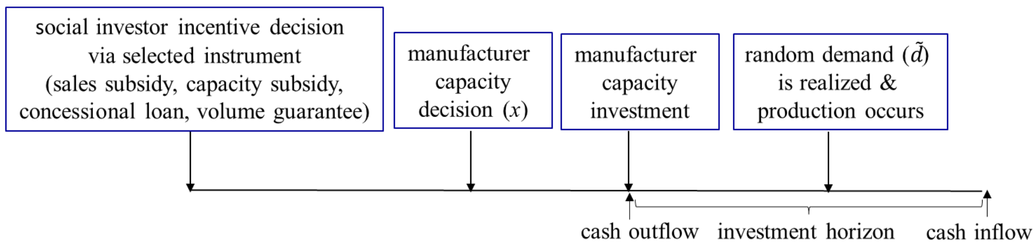

Figure 1. Timing of events.

The investment in capacity takes place at time zero. We denote the present value cost per unit of capacity by $\underline{c}_k$ and the fixed cost associated with building capacity by $\underline{c}_f$. Thus, at the beginning of the investment horizon, the cost to build capacity capable of producing and distributing up to x units during the investment horizon is then $\underline{c}_k x + \underline{c}_f$. We assume $\underline{c}_k > 0$ (i.e., cost is increasing in capacity, which is realistic and avoids the relatively trivial analysis that arises when $\underline{c}_k = 0$) and $\underline{c}_f > 0$. The marginal cost of capacity has a range of possible interpretations that depend on the specific setting. Recall the examples in Section 3.1. For the Jadelle contraceptive implant, $\underline{c}_k$ is the cost of building a new production line at an existing plant; for Hologic, $\underline{c}_k$ is the cost of building and installing test machines in an LMIC market; for bed nets, $\underline{c}_k$ is the cost of retrofitting existing production equipment and processes to accommodate a new insecticide.

$\tilde{d}$ = uncertain demand during investment horizon; pdf f, cdf F and complement $\bar{F} = 1 - F$, mean μ, variance σ^2
p = unit price paid by LMIC markets (end of investment horizon)
c_v = manufacturer variable cost of production (end of investment horizon)
i_2 = manufacturer cost of capital over investment horizon
i_1 = social investor loan rate over investment horizon
$\underline{c}_k$ = manufacturer cost per unit of capacity (start of investment horizon)
$c_k = (1+i_2)\underline{c}_k$ = manufacturer cost per unit of capacity (end of investment horizon)
$\underline{c}_f$ = manufacturer fixed cost to build capacity (start of investment horizon)
$c_f = (1+i_2)\underline{c}_f$ = manufacturer cost to invest in capacity (end of investment horizon)
υ = social value per unit of the health product
x = production/distribution capacity over investment horizon

Table 1. Model notation.

We express the firm's cost and revenue parameters at the end of the investment horizon (i.e., future value), and eliminate the underline on the cost parameters to denote the end-of-horizon values inflated by the manufacturer's cost of capital i_2, i.e., $c_k = (1+i_2)\underline{c}_k$ and $c_f = (1+i_2)\underline{c}_f$. We approximate the dynamics of financial flows over the investment horizon by assuming that random demand is realized at the end of the horizon. The manufacturer produces the minimum of demand and capacity at unit cost c_v. The upper limit of LMIC ability to pay is p. The model notation is summarized in Table 1.

To streamline notation, we normalize the residual value of the capacity investment at the end of the horizon to zero.[3] Therefore, the manufacturer's expected profit as a function of capacity x is

$$\pi(x) = (p - c_v) E \min\{\tilde{d}, x\} - c_k x - c_f \ . \tag{1}$$

Let $x_2^{\text{o}} = \arg\max_{x \geq 0} \{\pi(x)\}$. If $p - c_v \leq 0$, then $\pi(x)$ is decreasing in x and thus $x_2^{\text{o}} = 0$; otherwise, it follows from Proposition 1 that x_2^{o} satisfies

$$\bar{F}(x_2^{\text{o}}) = \min\{c_k / (p - c_v), 1\}. \tag{2}$$

Thus, the manufacturer's capacity decision and profit are

$$x_2^* = \begin{cases} 0, & \text{if } \pi(x_2^{\text{o}}) \leq 0 \\ x_2^{\text{o}}, & \text{if } \pi(x_2^{\text{o}}) > 0 \end{cases} \tag{3}$$

$$\pi^* = \begin{cases} 0, & \text{if } \pi(x_2^{\text{o}}) \leq 0 \\ \pi(x_2^{\text{o}}), & \text{if } \pi(x_2^{\text{o}}) > 0 \end{cases} . \tag{4}$$

Expression (3) reflects two possibilities. If $\pi(x_2^{\text{o}}) \leq 0$, then the variable profit at optimal capacity does not cover the fixed cost. In this case, the manufacturer has no incentive to invest in capacity (i.e., the manufacturer's participation constraint, $\pi(x) > 0$, is not satisfied). Alternatively, the manufacturer earns positive expected profit from building capacity (i.e., $\pi(x_2^{\text{o}}) > 0$), but the level is insufficient from the social investor's perspective. Both instances arise in practice. Referring to the examples in Section 3.1, $\pi(x_2^{\text{o}}) \leq 0$ for Jadelle and $\pi(x_2^{\text{o}}) > 0$ for Hologic and bed nets though investment in capacity is well below need from the social investors' perspective (see also Kazaz et al. 2021 for an underinvestment example).

Throughout the paper, we assume

$$x_2^* < x_1^* . \tag{5}$$

If (5) did not hold, then the social investor is not interested in incentivizing the manufacturer to invest in capacity.

[3] Parameter $\underline{c}_k$ can be interpreted as net of end-of-horizon salvage value, e.g., $\underline{c}_k = \kappa - s/(1 + i_2)$ where κ is the marginal cost of capacity incurred at time zero and s is the expected salvage value per unit of capacity at the end of the investment horizon.

4. Incentivizing Investment in Capacity

We consider four canonical instruments for incentivizing manufacturer investment in capacity.[4] The first two subsections describe the instruments. §4.1 describes two instruments from the literature and clarifies what is similar and what is new when applied in our setting. §4.2 describes two new instruments and presents properties of all four instruments. This is followed by two subsections, which explain a relationship between one canonical instrument and a generalization that arises in practice (§4.3), and present numerical results that illustrate regions of instrument dominance (§4.4).

4.1. Output-Unit and Input-Unit Instruments

As noted in Section 2, previous literature has studied the newsvendor order-quantity decision under a sales subsidy (payment to the newsvendor for each unit sold, or output-unit subsidy) and a purchase subsidy (payment to the newsvendor for each unit purchased, or input-unit subsidy), and find that a purchase subsidy dominates a sales subsidy. These instruments can be adapted to our setting where the decision is capacity instead of order quantity subject to a participation constraint (due to the presence of a fixed cost). Manufacturer expected profit as a function of capacity x and nonnegative instrument parameter y can be expressed as

$$\pi_s(x,y) = (p - c_v + y) E \min\{\tilde{d}, x\} - c_k x - c_f$$

$$\pi_k(x,y) = (p - c_v) E \min\{\tilde{d}, x\} - (1-y) c_k x - c_f$$

Instrument s is a *sales subsidy*; the manufacturer is paid y per unit of sales (output-unit subsidy). Instrument k is a *variable-capacity* subsidy; the manufacturer is paid fraction y of the variable cost of capacity (output-unit subsidy).[5]

The introduction of the manufacturer participation constraint that arises in our setting due to the fixed cost of capacity leads to more nuanced conclusions on instrument dominance. In the following, we first present results that characterize optimal decisions and profit given that the participation constraint is ignored. To simplify notation and focus on settings most relevant for practice, we assume that the social investor is interested in levels of capacity that are greater than the minimum possible demand, i.e., letting $\hat{x}_j(y)$ denote incentivized capacity as a function of y for instrument j, y satisfies

$$\hat{x}_j(y) > \min\{\tilde{d}\} \tag{6}$$

[4] These are instruments identified as the most critical set of alternatives by the Strategic Investment Fund team at the Gates Foundation.

[5] The manufacturer profit function under instrument k can be equivalently expressed as a payment on each unit of capacity (e.g., $(c_k - y)x$ in place of $(1 - y)c_k x$). We use y as a fraction because it is useful for comparative analysis of the instruments.

for all instruments (this assumption eliminates the need for a min operator in expressions for optimal decisions).

Proposition 2 (manufacturer profit at incentivized capacity). (*i*) *Ignoring the manufacturer's participation constraint*, *manufacturer's capacity decisions under each instrument are*

$$\hat{x}_s(y) = \arg\max_{x\geq 0}\{\pi_s(x,y)\} = \bar{F}^{-1}\left(\frac{c_k}{p-c_v+y}\right) \textit{ if } p - c_v + y > 0; \textit{ otherwise } \hat{x}_s(y) = 0 \tag{7}$$

$$\hat{x}_k(y) = \arg\max_{x\geq 0}\{\pi_k(x,y)\} = \bar{F}^{-1}\left(\frac{(1-y)c_k}{p-c_v}\right) \textit{ if } p - c_v > 0; \textit{ otherwise } \hat{x}_k(y) = 0. \tag{8}$$

(*ii*) *Ignoring the manufacturer's participation constraint, manufacturer's profits as a function of optimal capacity decision $x > x_2^{\mathrm{o}}$ and instrument parameter functions are*

$$\pi_s(x) = \pi_s(x, y_s(x)) = \left(\frac{c_k}{\bar{F}(x)}\right)E\left[\tilde{d} \mid \tilde{d} \leq x\right]F(x) - c_f$$

$$y_s(x) = \frac{c_k}{\bar{F}(x)} - (p - c_v)$$

$$\pi_k(x) = \pi_k(x, y_k(x)) = (p-c_v)E\left[\tilde{d} \mid \tilde{d} \leq x\right]F(x) - c_f \textit{ if } p - c_v > 0; \textit{ otherwise } \pi_k(x) = \varnothing \tag{9}$$

$$y_k(x) = 1 - \left(\frac{p-c_v}{c_k}\right)\bar{F}(x) \textit{ if } p - c_v > 0; \textit{ otherwise } y_k(x) = \varnothing \tag{10}$$

As shown in (8) – (10), instrument k is only viable if $p - c_v > 0$, i.e., no subsidy on capacity will incentivize investment in capacity if the manufacturer cannot make a positive profit on each unit sold.

With the above notation and results, we are now positioned to characterize investment in capacity for a given budget under each instrument while accounting for participation constraints. In other words, for a given cost to the social investor, what will be the manufacturer's capacity decision under each instrument?

Recall that incentivized capacity $x > x_2^*$ under instrument j is only achievable if $\pi_j(x) > 0$ (i.e., the participation constraint must be satisfied); if $\pi_j(x) \leq 0$, then the manufacturer will not invest in capacity. Also recall that x_2^{o} is the manufacturer's optimal investment in capacity (with no instrument) if the participation constraint is ignored. For instrument j, let

$$b_j^{\tau} = \begin{cases} 0, & \text{if } \pi\left(x_2^{\mathrm{o}}\right) > 0 \\ \pi_j\left(\pi_j^{-1}(0)\right) - \pi\left(\pi_j^{-1}(0)\right), & \text{if } \pi\left(x_2^{\mathrm{o}}\right) \leq 0 \end{cases}, \tag{11}$$

which is the minimum budget for which instrument j incentivizes investment in capacity, or *budget threshold*. We briefly explain the basis and meaning of (11). First, suppose that $\pi\left(x_2^{\mathrm{o}}\right) > 0$. This means that the manufacturer's participation constraint (without an instrument) is not binding, i.e., the manufacturer invests in capacity $x_2^* = x_2^{\text{o}} > 0$ without any subsidy. Thus, the budget threshold is zero. However, if $\pi\left(x_2^{\text{o}}\right) \le 0$, then the manufacturer will not invest without a subsidy. At $x = x_j^\tau := \pi_j^{-1}(0)$, manufacturer profit under instrument j is zero. Thus, instrument j is only viable for incentivizing capacity $x > x_j^\tau$. The corresponding cost to the social investor is the difference in manufacturer profits with, and without, the instrument,

$$b_j(x) = \pi_j(x) - \pi(x) \text{ for } x > x_j^\tau, \text{[6]} \tag{12}$$

which yields budget threshold $b_j^\tau = \pi_j\left(x_j^\tau\right) - \pi\left(x_j^\tau\right)$. Inverting (12) yields incentivized capacity for a given budget b, i.e., $x_j(b) = b_j^{-1}(b)$ for $b > b_j^\tau$.

Proposition 3 (instrument *s* versus *k*). (*i*) *Suppose* $p - c_v - c_f/\mu \le 0$. *Then*

$$0 < b_s^\tau < b_k^\tau = \infty$$

$$0 = x_k(b) < x_s(b) \textit{ for all } b > b_s^\tau.$$

(*ii*) *Suppose* $p - c_v - c_f/\mu > 0$ *and* $\pi\left(x_2^{\text{o}}\right) \le 0$. *Then*

$$0 < b_s^\tau < b_k^\tau < \infty$$

$$0 = x_k(b) < x_s(b) \textit{ for all } b \in \left(b_s^\tau, b_k^\tau\right]$$

$$0 < x_s(b) < x_k(b) \textit{ for all } b > b_k^\tau.$$

(*iii*) *Suppose* $p - c_v - c_f/\mu > 0$ *and* $\pi\left(x_2^{\text{o}}\right) > 0$. *Then*

$$0 = b_s^\tau = b_k^\tau$$

$$0 < x_2^* < x_s(b) < x_k(b) \textit{ for all } b > 0.$$

We offer several observations related to Proposition 3. First, suppose that $c_f = 0$, i.e., there is no participation constraint. Then, assuming gross margin is positive (i.e., $p - c_v > 0$), we see that a variable-capacity subsidy dominates a sales subsidy; if $p - c_v \le 0$, then s dominates k because k is not viable. This result is consistent with findings in the literature.

---

[6] We note that the social investor cost function for instrument j is only defined for x satisfying $\pi_j(x) > 0$ and $x > x_2^*$: $b_j(x) = \pi_j(x) - \pi(x)$ for $x > x_j^\tau := \min\{x : \pi_j(x) \ge 0\}$, e.g., if $x \le x_2^*$, then $b_j(x) = 0$. Note that $b_j(x)$ is strictly increasing in x. Therefore, the function can be inverted to obtain incentivized capacity as a function of budget, $x_j(b) = b_j^{-1}(b)$.

If $p - c_v > 0$, then the relationship among the two instruments becomes mixed in the presence of participation constraints (due to a fixed cost of the quantity decision). In particular, the sales subsidy can dominate the variable-capacity subsidy at low budget levels (e.g., budget levels where k is not viable). The intuition is as follows. As noted earlier, in the absence of participation constraints, the manufacturer's capacity decision is more sensitive to an input-unit subsidy than an output-unit subsidy; an input-unit subsidy reduces both the underage and overage cost in the newsvendor model whereas an output-unit subsidy only reduces underage cost. The main consequence from a social investor's perspective is the lower leverage from an output-unit subsidy means "leaving money on the table" for a given capacity compared to the input-unit subsidy. However, in the presence of participation constraints, the apparent negative of extra cost can become a positive because it can cover the manufacturer's fixed cost thereby incentivizing manufacturer participation at a lower budget than what is possible via instrument k.

4.2. Total-Capacity and Volume Guarantee Instruments

In this section, we first describe two new instruments. Then we present results that characterize settings under which a particular instrument is preferred.

$$\pi_l(x,y) = (p - c_v)E\min\{\tilde{d}, x\} - (1-y)(c_k x + c_f) \tag{13}$$

$$\pi_q(x,y) = (p - c_v)E\min\{\max\{\tilde{d}, y\}, x\} - c_k x - c_f \,. \tag{14}$$

Instrument l is a *total-capacity* subsidy; the manufacturer is paid fraction y of the total cost of capacity. This instrument captures the structure of manufacturer profit under a concessional loan. We present details of a concessional loan in Section 4.3. If $c_f = 0$, then instruments k and l are equivalent. We assume $c_f > 0$ throughout the remainder of the paper.[7] Instrument q is a *volume-guarantee* subsidy because the manufacturer is assured sales volume of at least y units; given realized demand d, the manufacturer is paid $(p - c_v)(y - \min\{d, x\})^+$. Instrument q is a risk-shifting instrument—it shifts risk associated with low demand (and consequent low profit) from the manufacturer to the social investor. Note that if $y = 0$, then the profit functions reduce to (1), i.e., $\pi_j(x, 0) = \pi(x)$ for $j \in \{s, k, l, q\}$.

Proposition 4 extends Proposition 2 for instruments l and q. Let

$$\bar{F}_{\max\{\tilde{d},y\}}(x) = \Pr\left(\max\{\tilde{d}, y\} > x\right).$$

Proposition 4 (manufacturer profit at incentivized capacity). (*i*) *Ignoring the manufacturer's participation constraint, manufacturer's capacity decisions under instruments l and q are*

$$\hat{x}_l(y) = \arg\max_{x\geq 0}\{\pi_l(x,y)\} = \bar{F}^{-1}\left(\frac{(1-y)c_k}{p - c_v}\right) \text{ if } p - c_v > 0; \text{ otherwise } \hat{x}_l(y) = 0 \tag{15}$$

---

[7] The consequence of this assumption is that some strict inequalities in our results become non-strict due to the equivalence of instruments k and l when $c_f = 0$.

$$\hat{x}_q(y) = \arg\max_{x \geq 0} \{\pi_q(x, y)\} = \bar{F}^{-1}_{\max\{\tilde{d}, y\}}\left(\frac{c_k}{p - c_v}\right) \text{ if } p - c_v - c_k > 0; \text{ otherwise } \hat{x}_q(y) = 0. \tag{16}$$

(*ii*) *Ignoring the manufacturer's participation constraint, manufacturer's profits as a function of optimal capacity decision* $x > x_2^{\text{o}}$ *and instrument parameter functions are*

$$\pi_l(x) = (p - c_v)E\left[\tilde{d} \mid \tilde{d} \leq x\right]F(x) - \left(\frac{\bar{F}(x)}{c_k / (p - c_v)}\right)c_f \text{ if } p - c_v > 0; \text{ otherwise } \pi_k(x) = \varnothing$$

$$y_l(x) = 1 - \left(\frac{p - c_v}{c_k}\right)\bar{F}(x) \text{ if } p - c_v > 0; \text{ otherwise } y_l(x) = \varnothing$$

$$\pi_q(x) = (p - c_v - c_k)x - c_f \text{ if } p - c_v - c_k > 0; \text{ otherwise } \pi_q(x) = \varnothing$$

$$y_q(x) = x \text{ if } p - c_v - c_k > 0; \text{ otherwise } y_q(x) = \varnothing.$$

From $\bar{F}(x_2^{\text{o}}) = \min\{c_k/(p - c_v), 1\}$, it follows that $\bar{F}(x) < c_k/(p - c_v)$ for $x > x_2^{\text{o}}$. Therefore, given that $p - c_v > 0$, it follows from propositions 2 and 4 that

$$\pi_k(x) < \pi_l(x) < \pi_s(x) \text{ for all } x > x_2^{\text{o}}. \tag{17}$$

As we clarify later, the inequality between $\pi_q(x)$ and $\pi_j(x)$ for $j \in \{s, k, l\}$ can be in either direction.

Proposition 5 characterizes relative differences in incentivized capacity among the four instruments in a setting where the manufacturer will not invest in capacity without an instrument, i.e., $\pi(x_2^{\text{o}}) \leq 0$, which implies $x_2^* = 0$. Proposition 6 presents result for the opposite setting, i.e., $\pi(x_2^{\text{o}}) > 0$ and $x_2^* = x_2^{\text{o}} > 0$.

Proposition 5 (no capacity investment without an instrument). (*i*) *Suppose* $p - c_v \leq 0$. *Then*

$$0 < b_s^\tau < b_k^\tau = b_l^\tau = b_q^\tau = \infty$$

$$0 = x_k(b) = x_l(b) = x_q(b) < x_s(b) \text{ for all } b > b_s^\tau.$$

(*ii*) *Suppose* $p - c_v - \max\{c_k, c_f/\mu\} \leq 0 < p - c_v$. *Then*

$$0 < b_s^\tau = b_l^\tau < b_k^\tau = b_q^\tau = \infty$$

$$0 = x_k(b) = x_q(b) < x_s(b) < x_l(b) \text{ for all } b > b_s^\tau = b_l^\tau.$$

(*iii*) *Suppose* $p - c_v - c_k \leq 0 < p - c_v - c_f/\mu$. *Then*

$$0 < b_s^\tau = b_l^\tau < b_k^\tau < b_q^\tau = \infty$$

$$0 = x_k(b) = x_q(b) < x_s(b) < x_l(b) \text{ for all } b \in \left(b_s^\tau, b_k^\tau\right] = \left(b_l^\tau, b_k^\tau\right]$$

$$0 = x_q(b) < x_s(b) < x_l(b) < x_k(b) \text{ for all } b > b_k^\tau.$$

(iv) *Suppose* $p - c_v - c_f/\mu \leq 0 < p - c_v - c_k$. *Then*

$$0 < b_s^\tau = b_l^\tau < \infty \text{ and } 0 < b_q^\tau < b_k^\tau = \infty$$

$$0 = x_k(b) \text{ for all } b$$

$$0 < x_s(b) < x_l(b) \text{ for all } b > b_s^\tau = b_l^\tau$$

(v) *Suppose* $0 < p - c_v - \max\{c_k, c_f/\mu\}$. *Then*

$$0 < b_s^\tau = b_l^\tau < b_k^\tau < \infty \text{ and } 0 < b_q^\tau < \infty$$

$$0 = x_k(b) = x_q(b) < x_s(b) < x_l(b) \text{ for all } b \in \left(b_s^\tau, b_k^\tau\right] = \left(b_l^\tau, b_k^\tau\right]$$

$$0 < x_s(b) < x_l(b) < x_k(b) \text{ for all } b > b_k^\tau.$$

Proposition 6 (capacity investment occurs without an instrument).

$$0 < p - c_v - \max\{c_k, c_f/\mu\}$$

$$0 = b_s^\tau = b_k^\tau = b_l^\tau = b_q^\tau$$

$$0 < x_2^* < x_q(b) < x_s(b) < x_s(b) < x_k(b) \text{ for all } 0 < b \leq (p - c_v) E\left(x_2^* - \tilde{d}\right)^+$$

$$0 < x_2^* < x_s(b) < x_l(b) < x_k(b) \text{ for all } b > 0.$$

Let us summarize the main conclusions from the propositions and the underlying intuition. If the manufacturer is willing to invest in capacity without a subsidy instrument, then instrument k consistently dominates instruments s and l (Proposition 3). Furthermore, a volume guarantee at capacity $x_2^* > 0$ incurs an expected cost to the social investor of $(p - c_v)E\left(x_2^* - \tilde{d}\right)^+$, which implies $x_q(b) = x_2^*$ for all $b \leq (p - c_v)E\left(x_2^* - \tilde{d}\right)^+$, e.g., a necessary condition for $x_q(b) > x_k(b)$ is $b > (p - c_v)E\left(x_2^* - \tilde{d}\right)^+$.

As noted above, the dominance of an input subsidy over an output subsidy derives from a difference in leverage within a newsvendor structure. An increase in the price subsidy increases the shortage cost per unit, $p - c_v - c_k + y$, whereas an increase in the variable cost subsidy increases the shortage cost per unit, $p - c_v - (1 - y)c_k$ and simultaneously decreases the excess cost per unit, $(1 - y)c_k$. Instrument k dominates instrument l in this setting because decreasing the fixed cost (via l) does not affect the marginal value of capacity for the manufacturer, e.g., compare (8) with (15).

The picture notably changes in settings where the manufacturer is not willing to invest in capacity without a subsidy. In this setting, instrument s is the only viable instrument when ability to pay p does not exceed the variable production cost c_v. As ability to pay net of variable production cost increases above zero but remains below the variable capacity cost (c_k) and below the fixed cost of capacity per unit of forecasted demand (c_f/μ), then instrument l is viable and dominates instrument s. The dominance of instrument l relative to s derives from the greater leverage of a subsidy on an input unit over an output unit, as discussed above. Interestingly, while both instruments l and s are not viable at low budget levels, the instruments share the same budget threshold. This is due to the structural similarity of profit under the two instruments, e.g., instrument s increases revenue by factor $(p - c_v + y)/(p - c_v)$, which when factored out yields profit as the product $(p - c_v + y)/(p - c_v)$ and a profit expression with capacity cost reduced by factor $(p - c_v)/(p - c_v + y)$, which conforms to the structure of profit under instrument l.

As ability to pay increases above the sum of variable production cost and the fixed cost of capacity per unit of forecasted demand, instrument k becomes viable when budget is high enough, i.e., above the budget threshold b_k^τ that is higher than the threshold for instruments l and s. For budgets above b_k^τ, instrument k dominates l and s for reasons explained above.

Instrument q cannot incentivize investment in capacity when ability to pay does not exceed the variable cost of production and capacity. Under a volume guarantee, the manufacturer will either not build capacity or will set capacity to match the guaranteed volume. Consequently, instrument q is only viable at levels of capacity that exceed breakeven volume $c_f/(p - c_v - c_k)$. Clearly, breakeven cannot be achieved if $p - c_v - c_k \leq 0$. Given that $p - c_v - c_k > 0$, the inequality among the instrument q budget threshold and the other budget thresholds, and the inequality among the instrument q incentivized capacity function and the other capacity functions can go in either direction. This can be explained by a fundamental difference in in the structure of the manufacturer's capacity decision that is evident by comparing (16) with (7), (8), and (15) (e.g., decisions are tied to different probability distributions). Table 2 interprets the results from the propositions to identify settings where a particular instrument offers the highest investment in capacity for a given budget. These settings are distinguished by up to three dimensions: (1) ability to pay from very low to high, (2) capacity cost spanning high variable, high fixed, moderate variable and fixed, (3) budget from low to high.

Capacity investment is unprofitable for the manufacturer	Dominant instrument
A. Very low ability to pay: $p - c_v \leq 0$	sales subsidy
B. Low ability to pay: $p - c_v - \max\{c_k, c_f/\mu\} \leq 0 < p - c_v$	total-capacity subsidy
C. Moderate ability to pay, high variable capacity cost: $p - c_v - c_k \leq 0 < p - c_v - c_f/\mu$	low budget: total-capacity subsidy high budget: variable-capacity subsidy
D. Moderate ability to pay, high fixed capacity cost: $p - c_v - c_f/\mu \leq 0 < p - c_v - c_k$	total-capacity subsidy or volume guarantee
E. Moderate ability to pay, moderate capacity costs: $0 < p - c_v - \max\{c_k, c_f/\mu\}$ and $\pi(x_2^o) \leq 0$	low budget: total-capacity subsidy or volume guarantee high budget: variable-capacity subsidy or volume guarantee
Capacity investment is profitable for the manufacturer $0 < p - c_v - \max\{c_k, c_f/\mu\}$ and $\pi(x_2^o) > 0$	**Dominant instrument**
F. High ability to pay with low budget	variable-capacity subsidy
G. High ability to pay with high budget	variable-capacity subsidy or volume guarantee

Table 2. Regions of instrument preference.

4.3. Concessional Loan

A concessional loan refers to a low-interest loan to the manufacturer, i.e., a loan for all or a portion of the manufacturer's investment in capacity. This instrument may be employed when the social investor has sufficient cash reserves to provide low-interest loans at rate $i_1 < i_2$ where i_1 is typically linked to the social investor's return on capital. The social investor does not seek returns on its philanthropic capital (as per Internal Revenue Service). The nonprofit social investor's monies are typically structured as follows. The donated capital to the nonprofit is kept in an investment trust. The trust manages the capital as investment assets by investing them in a portfolio of return-yielding instruments. As the nonprofit needs more money for its programs (including the types of investments discussed in this paper), the trust fund must liquidate part of the portfolio and/or use earnings from the trust to transfer to the nonprofit's programming arm. These structures are slightly different in the UK, the US, and other parts of the world. However, in each case the nonprofit social investor forgoes earnings from its invested assets when it makes a large investment and forgoes the opportunity to invest these funds in other activities related to the nonprofit/charitable goals. We have validated that this approach is what drives internal decision-making regarding the cost of capital at the major groups that make such investments.

A concessional loan lowers the manufacturer's cost of capital. The savings to the manufacturer from a low-interest loan at rate i_1 for fraction z of the total investment is structurally equivalent to a total-capacity subsidy. To clarify this point, recall that the end-of-horizon cost to the manufacturer of an investment in capacity is $c_k x + c_f = (1+i_2)(\underline{c}_k x + \underline{c}_f)$. If the manufacturer receives a concessional loan on fraction z of the investment, the end-of-horizon cost of the investment is

$$(1+i_1)z(\underline{c}_k x + \underline{c}_f) + (1+i_2)(1-z)(\underline{c}_k x + \underline{c}_f) = \left(1 - \frac{z(i_2 - i_1)}{1+i_2}\right)(c_k x + c_f)$$

and the manufacturer profit is given in (13) with

$$y = \frac{z(i_2 - i_1)}{1+i_2}.$$

While a concessional loan is equivalent to a total-capacity subsidy from the manufacturer's perspective, this is not necessarily the case for the social investor. We introduce parameter $\iota \geq 0$ to capture the social investor's cost of a concessional loan as a fraction of the manufacturer's savings; that is, the manufacturer savings for a concessional loan (and total-capacity subsidy) given capacity x is

$$y(c_k x + c_f)$$

and the social investor cost is

$$\iota \times y(c_k x + c_f).$$

Clearly, if $\iota = 1$, then there is no difference between a total-capacity subsidy and a concessional loan from the social investor's perspective. Differences arise when $\iota \neq 1$, wherein a concessional loan is less (more) costly than a total-capacity subsidy when $\iota < 1$ ($\iota > 1$).

The effects of parameter ι on the results in propositions 3 and 4 are relatively straightforward;[8] instrument l becomes more ($\iota > 1$) or less expensive ($\iota < 1$) as ι deviates from 1. Let $\overline{x}_l(b)$ and $\overline{b}_l^{\tau}$ denote the capacity function and budget threshold for a concessional loan. In effect, parameter b in function $x_l(\bullet)$ is replaced by ιb in function $\overline{x}_l(\bullet)$, or equivalently, b in $\overline{x}_l(\bullet)$ is replaced by b/ι in $x_l(\bullet)$.[9]

Corollary 1 (instrument is not needed for capacity investment). *Suppose that* $x_2^* > 0$. *Then*

$$\overline{b}_l^{\tau} = b_l^{\tau} = 0$$

$$\overline{x}_l(b) = x_l(b/\iota) \text{ for all } b > 0$$

[8] Our communication with Strategic Investment Fund team at the Gates Foundation indicates that excluding fixed costs of building capacity (that are generally significant) has not historically been considered, in part because it adds complexity to the instrument. However, it is straightforward to generalize the variable-capacity subsidy to include parameter ι, which yields a model of a concessional loan for a fraction the variable cost of capacity (i.e., fixed cost is excluded from the loan).

[9] For the extreme of $\iota = 0$, the cost to the social investor is zero, though manufacturer profit and subsidy parameter functions for instrument l in Proposition 2 continue to apply.

$x_k(b) > \hat{x}_l(b)$ *for all* $b > 0$ *if* $\iota \geq 1$

$\overline{x}_l(b) > x_s(b)$ *for all* $b > 0$ *if* $\iota \leq 1$.

Corollary 2 (instrument is needed to increase capacity investment). *Suppose that* $x_2^* = 0$. (*i*) *Suppose that* $p - c_v \leq 0$. *Then* $\overline{x}_l(b) = 0$ *for all* $b > 0$. (*ii*) *Suppose that* $p - c_v > 0$. *Then* $\overline{b}_l^\tau = \iota b_l^\tau$ *and* $\overline{x}_l(b) = x_l(b/\iota)$ *for all* $b > \overline{b}_l^\tau$.

4.4. Regions of Instrument Dominance

Our earlier propositions and Table 2 have collectively examined how the preferred instrument is affected by budget b, ability to pay p, and nature of variable and fixed capacity costs. In this section, we investigate the regions of instrument dominance in the combination of budget and ability pay (i.e., in (p, b) space), and we examine how differences in the relationship between variable and fixed capacity costs affect the preferred instrument.

We assume throughout the remainder of the paper that forecast error is normally distributed. The normal distribution is often a reasonable approximation for forecast error due to the Central Limit Theorem (e.g., forecast error is the aggregation of noise terms across many buyers).

We limit consideration of instrument l as described in Section 4.2, i.e., we set $\iota = 1$, as the directional effect of an increase or decrease in ι on the power of instrument l is clear. Table 2 in Section 4.1 identifies regimes where the dominant instrument is clear (i.e., A, B, and F in Table 2) and other regimes where the dominant instrument is ambiguous (i.e., C, D, E, and G in Table 2). In this section, we illustrate characteristics of settings where a particular instrument dominates. Recall that a sales subsidy dominates all instruments if and only if ability to pay is not more than variable cost of production (i.e., $p \leq c_v$). Consequently, we limit our numerical illustrations to cases where $p > c_v$.

The social optimal capacity is $x_1^* = F^{-1}(\alpha_1^*)$ where $\alpha_1^* = 1 - c_k/(\upsilon - c_v)$ is the social optimal service level. Figures 2 and 3 identify the instrument in (p, b) space that achieves the highest investment in capacity up to social optimal x_1^* (i.e., social welfare decreases as capacity increases beyond x_1^*) where x_1^* corresponds to social optimal service level of $\alpha_1^* = 0.90$. Figure 2 illustrates regions of instrument dominance for an example with moderate fixed capacity cost and high variable capacity cost, $c_f/\mu < c_k$ (e.g., instrument k becomes viable before instrument q as ability to pay increases). Figure 3 illustrates regions of instrument dominance for an example with moderate variable capacity cost and high fixed capacity cost, $c_k < c_f/\mu$ (e.g., instrument q becomes viable before instrument k as ability to pay increases). The demand forecast is $\mu = 100$ for both figures. Other parameter values are identified under each figure.

To illustrate the interpretation of results in the figures, consider Figure 2 at p = 16. Instrument l incentivizes the highest investment in capacity for a budget between 960 (= budget threshold for l) and 1170. Instrument k incentivizes greater investment in capacity for a budget between 1171 (= budget threshold for k) and 1565. The social optimal capacity is achieved with instrument k and b = 1565 (i.e., dominance regions in the plots are identified up to social optimal capacity).

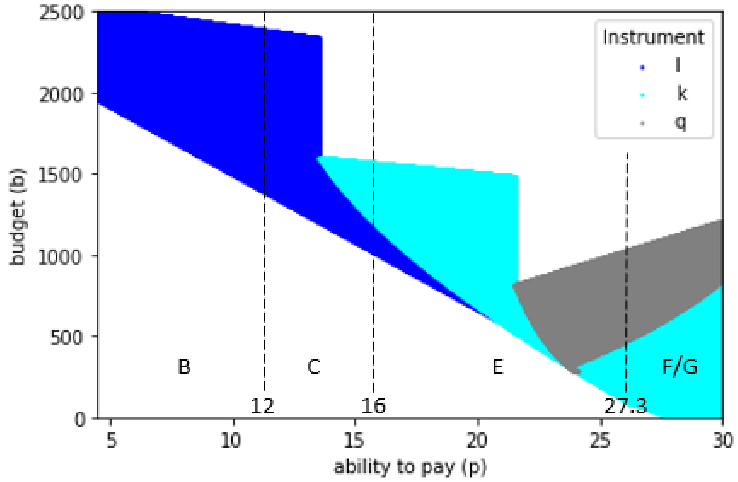

Figure 2. Regions in (p, b) space where instrument l (blue), q (gray), or k (cyan) dominate for an example where $c_k > c_f/\mu$ (regimes B, C, E, F, G). The horizontal axis is ability to pay (p) and the vertical axis is budget (b). The data for the example: μ = 100, σ = 35, c_v = 4, c_k = 12, c_f = 800. The vertical lines delimit the regimes by ability to pay, e.g., the vertical line that appears at p = 27.3 is the minimum price at which the manufacturer invests in capacity without a subsidy (at service level = 48.4%).

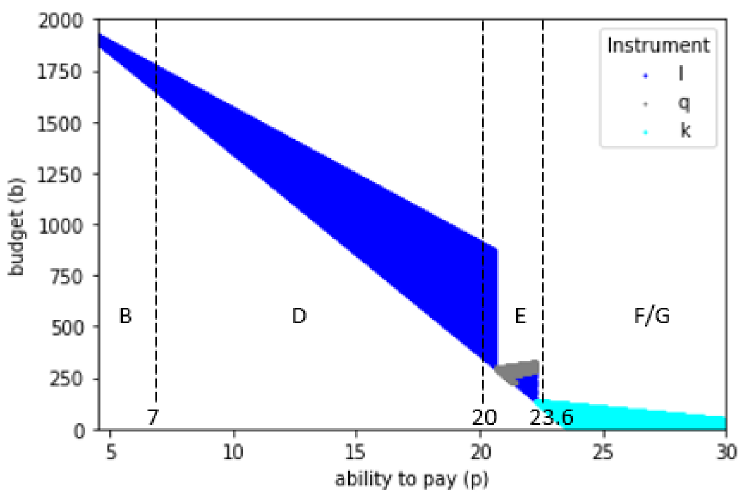

Figure 3. Regions in (p, b) space where instrument l (blue), q (gray), or k (cyan) dominate for an example where $c_k > c_f/\mu$ (regimes B, C, E, F, G). The horizontal axis is ability to pay (p) and the vertical axis is budget (b). The data for the example: μ = 100, σ = 13.5, c_v = 3, c_k = 4, c_f = 1600. The vertical lines delimit the regimes by ability to pay, e.g., the vertical line that appears at p = 23.6 is the minimum price at which the manufacturer invests in capacity without a subsidy (at service level = 84.8%).

We briefly sketch the process for creating the figures. The process has subtlety because the optimal budget can decrease in capacity due to instrument thresholds, and consequently, optimal incentivized capacity can exhibit jumps at budget points where the optimal instrument changes.[10] For each ability-to-pay, we solve

$$j^*(x) = \underset{j\in\{k,l,q\}}{\arg\min}\{b_j(x)\} \text{ for all } x \le x_1^*, \tag{18}$$

then sort $b_{j^*(x)}$ from smallest-to-largest and remove all $b_{j^*(x)}$ that satisfy $b_{j^*(x)} \ge b_{j^*(x')}$ for some $x' > x$ (because instrument $j^*(x)$ is dominated by instrument $j^*(x')$). Figures 2 and 3 report $b_{j^*(x)}$ at values of ability to pay that range between 4.5 and 30, which span regimes B through G in Table 2.

We present two related figures to support interpretations of figures 2 and 3. Figure 4 displays the threshold curves—service level $\alpha_j^\tau = F(x_j^\tau)$ (left plot) and budget b_j^τ (left plot)—for the calibration used in Figure 2. Figure 5 displays the curves for the calibration used in Figure 3. The left plots also identify the manufacturer's optimal service level without a subsidy instrument (i.e., manufacturer's newsvendor ratio).

The capacity threshold curves in the left plots of figures 4 and 5 illustrate the structural difference between volume guarantee (instrument q) and the other instruments. Excluding instrument q, capacity threshold curves of other instruments all intersect at ability to pay p where the manufacturer is willing to invest in capacity without a subsidy. It should be noted here that the volume guarantee (q) can incentivize much lower investments in capacity at slightly lower levels of ability to pay. This distinctive property allows q to become the dominant instrument at low budget levels in some settings (as illustrated in figures 2 and 3 in regime E). Finally, we note that differences in budget thresholds that abide by inequalities appearing in Proposition 5 can be significant (e.g., Figure 4 right plot).

[10] For example, for the Figure 2 calibration with $p = 14$, the capacity threshold for instrument k is $x_k^\tau = 140$. The least-cost instrument to incentivize capacity 139 is l with budget \$2205, whereas the least-cost instrument to incentivize capacity 140 is k with budget \$1501. In this example, instrument k dominates instrument l at budget \$1501 (assuming social optimal capacity is 140 or more), and there is a jump in incentivized capacity when budget increases from 1500 to 1501; l is the optimal instrument at $b = 1500$ yielding capacity $x_l(1500) = 112$, and k is the optimal instrument at $b = 1501$ yielding capacity $x_k(1501) = 140$.

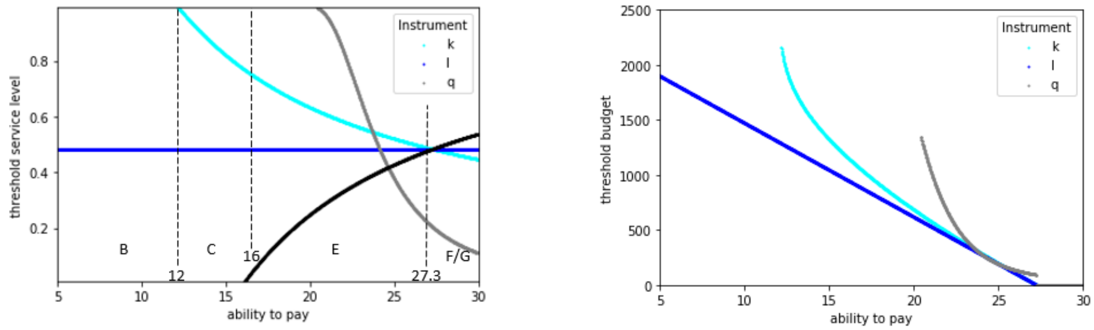

Figure 4. The left plot shows the threshold service level for instrument *l* (blue), *q* (gray), and *k* (cyan). The black line shows the manufacturer's newsvendor ratio. The right plot shows the threshold budget for the three instruments. The threshold service level for *l* (and *s*) is fixed at 48.4%, which is the service level where the three curves intersect at ability to pay = $27.3. The calibration is the same as Figure 2.

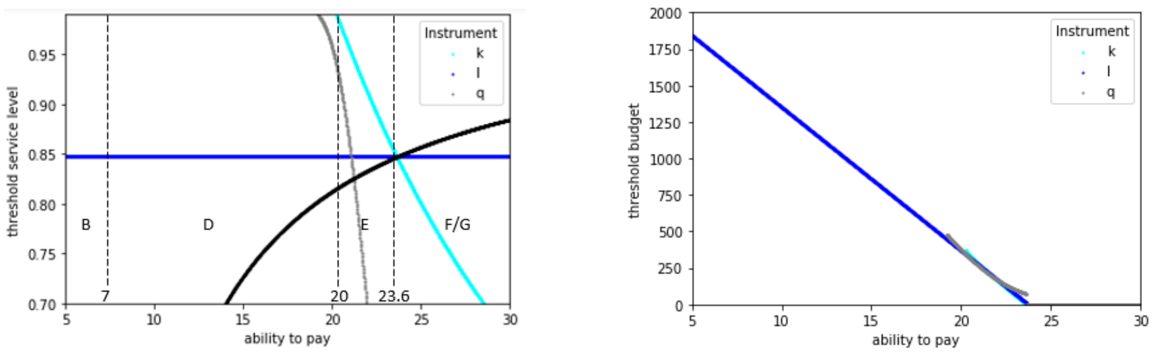

Figure 5. The left plot shows the threshold service level for instrument *l* (blue), *q* (gray), and *k* (cyan). The black line shows the manufacturer's newsvendor ratio. The right plot shows the threshold budget for the three instruments. The threshold service level for *l* (and *s*) is fixed at 84.8%, which is the service level where the three curves intersect at ability to pay = $23.6. The calibration is the same as Figure 3.

Figures 2 – 5 illustrate that regions of instrument dominance are not simple. However, at a high level, there is basic pattern that emerges that depends on whether variable cost of capacity (c_k) is low or high relative to fixed cost of capacity per forecasted unit. In all cases (as noted in Table 2), instrument *s* dominates at very low *p* (not included in the figures) and *l* dominates at low *p*. At moderate *p*, instrument *l* tends to dominate when budget is low. At high budget levels, the picture is more complex with instrument *k* tending to dominate when the ratio of fixed-to-variable capacity cost ($(c_f/\mu)/c_k$) is low and instrument *q* tending to dominate otherwise. At high *p* (e.g., to the point where the manufacturer is willing to invest in capacity without incentives), instrument *k* tends to dominate at low budget but may be supplanted by instrument *q* when the budget is high. These high-level lessons are summarized in the form of a strategy grid in Figure 6.

	low budget	high budget
high ability to pay	variable-capacity subsidy tends to dominate	variable-capacity subsidy or volume guarantee tends to dominate
moderate ability to pay	total-capacity subsidy tends to dominate	high $(c_f/\mu)/c_k$: volume guarantee tends to dominate low $(c_f/\mu)/c_k$: variable-capacity subsidy tends to dominate
low ability to pay	total-capacity subsidy dominates	total-capacity subsidy dominates
very low ability to pay	sales subsidy dominates	sales subsidy dominates

Figure 6. Strategy grid illustrating conditions under which a particular instrument dominates or is likely to dominate the other instruments.

5. Numerical Illustrations of Social Investor Instruments

In this section, we illustrate how our model applies on three real-world cases. Our goal in this section is not to develop new technical results beyond what is developed in sections 3 and 4. Our numerical analysis is also not an attempt to evaluate or second-guess actions for a given case. Rather, this section presents how our model helps decision makers identify the best instrument to incentivize capacity to serve socially desirable output of health products.

For our calibrations, we use data from the markets for which an instrument has been used in the past. We combine publicly available information on market parameters with estimates from unpublished studies and private sources.[11] As such, we mask cost-related parameters by normalizing the marginal cost of capacity c_k to 1, and proportionally adjusting the remaining parameters. The three products analyzed in our numerical illustration include a sufficient degree of heterogeneity in the parameters associated with the market and manufacturer's cost terms. They also represent a range of different health products. We emphasize that our purpose is to illustrate the application of our model. We seek reasonable estimates of values based on available data and communications with those who have first-hand knowledge. The values of the parameters in our calibration are listed in Table 3.

[11] Sources for estimating parameters varied by product. We summarize the nature of sources here, some of which are internal and confidential (i.e., estimated by the social investor). Demand forecast parameter estimates: public announcements of volumes, historical data for an earlier generation of a product. Manufacturer cost structure: published studies of COGS analysis, annual reports (for cost of capital). Ability to pay (WTP): internal studies, published studies of WTP on comparable products, historical prices paid for comparable product. Social value per unit: cost effectiveness studies published for comparable product, internal analysis.

The values of i_2 and i_1 in Table 3 warrant explanation. Recall that the length of a period in our model corresponds to the length of the investment horizon. Our calibration of i_2 uses a mapping from an annual rate to a rate for the investment horizon. We summarize our mapping here and refer the reader to Appendix C in the appendix for additional details. Let T denote the duration of the investment horizon in years. Recall that π_2 in our model is manufacturer profit at time T, and that i_2 represents the manufacturer's cost of capital over the investment horizon. For a given annul cost of capital, denoted r, one alternative is to apply annual compounding over the investment horizon to obtain

$$i_2 = (1 + r)^T - 1. \tag{19}$$

This alternative is exact if the investment occurs at time zero and all payoffs occur at time T. While there is a delay between initial cash outflow and the beginning of cash inflows due to sales (e.g., time to build/deploy capacity), a delay of no cash inflow until time T is extreme for the cases we consider, e.g., (19) may significantly overstate the value of i_2. Instead, we use a mapping from r to i_2 (and similarly to i_1) that assumes a payoff at the end of each year equal to fraction $1/T$ of the total cash inflow over the investment horizon, i.e.,

$$i_2 = \left(\frac{rT}{(1+r)^T - 1} \right) (1+r)^T - 1 . \tag{20}$$

The term in the first parentheses is the growth rate over time T with no compounding divided by the growth rate over time T with annual compounding. We note that any calibration of parameter values is not exact, and there may be alternative reasonable approaches to estimate i_2 from an annual rate. However, our discussions with those who have first-hand knowledge worked on the three cases indicate that (20) is reasonable for the numerical illustrations.

Parameter	**Product A**	**Product B**	**Product C**
Investment horizon (T)	6 years	3 years	3 years
Demand forecast over investment horizon (μ)	66m	54.8m	134.1m
Forecast error (σ)	9.2m	7.4m	50.1m
LMIC ability to pay (p)	$9.88	$6.00	$7.00
Manufacturer variable cost (c_v)	$7.00	$4.00	$4.00
Manufacturer unit capacity cost ($\underline{c}_k$)	$1.00	$1.00	$1.00
Manufacturer fixed cost ($\underline{c}_f$)	$80m	$12.5m	$45m
Manufacturer cost of capital over the investment horizon (i_2)	25.5%	53.7%	16.4%
Social investor concessional loan rate over the investment horizon (i_1)	18.2%	10.2%	2.0%
Ratio of social investor cost to manufacturer savings (ι)	1.00	0.25	0.50
Social value per unit of the health product (υ)	$63.08	$10.00	$10.00

Table 3. Parameters for the three products used for numerical illustration (m = 1 million).

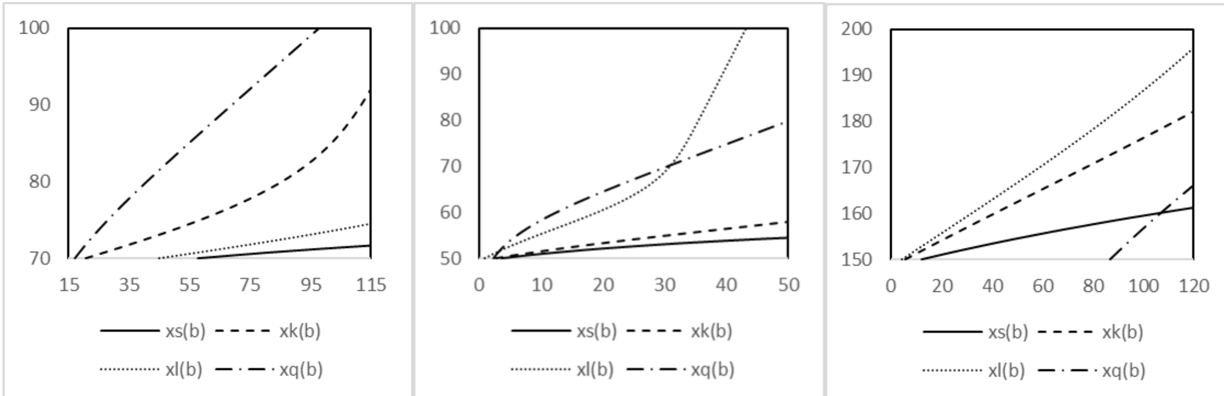

Figure 7. Incentivized capacity (vertical axis) as a function of budget (horizontal axis) for the four instruments – sales subsidy ($x_s(b)$), capacity subsidy ($x_k(b)$), concessional loan ($x_l(b)$), volume guarantee ($x_q(b)$) for Product A (left plot), Product B (middle plot), and Product C (right plot).

Ignoring the participation constraint, the optimal capacities for products A, B and C are x_2^o = 68m, 49m, 148m, respectively, which compare with the respective socially optimal capacities of x_1^* = 84m, 60m, and 177m. For product A, the manufacturer will not invest in capacity without a subsidy (x_2^* = 0). The manufacturer will invest in capacity without a subsidy for products B (x_2^* = 49) and C (x_2^* =148). Figure 7 shows some diversity in the dominant instrument over the range of budgets reported: a volume guarantee for product A, a volume guarantee for product B at low budgets and a concessional loan at high budgets, and a concessional loan for product C.

6. Summary and Implications for Social Investors

We present a framework to help social investors evaluate instruments for incentivizing a manufacturer to invest in production and distribution capacity. We focus on LMIC markets that are less desirable by manufacturers to invest in sufficient capacity. The framework illuminates how relevant factors interact to influence the relative attractiveness of different instruments. The same framework is also a springboard to numerically evaluate instrument costs. Barriers to manufacturer investment stem from a combination of lower margins and higher risk, compared to developed-country markets. Margins are lower because of a lower ability to pay, and they may be reduced because of higher costs of distribution due to less developed infrastructure. Risk is higher because of greater market uncertainty due to limited data on health product needs and opacity around government approvals of reimbursement lists. Collectively, these challenges elevate the importance of our framework developed specifically for LMIC markets in incentivizing capacity investments for global health products.

We consider four basic types of subsidy instruments in our framework–-three that target the barrier of low margin through payments to increase revenue or reduce cost, and one that targets the barrier of high risk through a volume guarantee. A sales subsidy pays the manufacturer an amount for each unit sold, a variable-capacity subsidy pays the manufacturer an amount for each unit of capacity built, and a total-capacity subsidy (and its close relative: concessional loan) pays the manufacturer an amount for each dollar spent (fixed and variable) on capacity.

Each instrument has its advantages that raise it to the preferred choice in some settings. Conditions under which a particular instrument is preferred are straightforward in some cases and are nuanced in other cases. A sales subsidy is the only viable instrument when ability to pay is very low. It is the most generous instrument from the perspective of the manufacturer, which is what causes the sales subsidy to be dominated by other instruments when ability to pay exceeds marginal production cost. The total-capacity subsidy is preferred when ability to pay exceeds marginal production but is below two measures of unit cost:[12] (1) the sum of marginal production and marginal capacity cost, (2) the sum of marginal production cost and fixed cost capacity per unit of forecasted demand. When condition (1) holds, a volume guarantee is not viable because each guaranteed unit of sales returns a loss to the manufacturer. When condition (2) holds, a variable-capacity subsidy is not viable because it is not generous enough to satisfy the manufacturer's participation constraint, i.e., manufacturer gross profit (excluding fixed capacity cost) at any incentivized capacity does not exceed the fixed cost of capacity.

When condition (1) holds, but condition (2) does not, then a total-capacity subsidy is preferred when budget is low, while the variable-capacity subsidy is preferred at higher budgets. Alternatively, when condition (1) holds, but condition (2) does not, then the dominant instrument is either a total-capacity subsidy or a volume guarantee with the preferred instrument dependent on parameter values and budget. If neither condition holds, then the preferred instrument among the three depends on parameter values and budget with one exception—if the manufacturer is willing to invest in capacity (but at a level below is desired by the social investor), then the total-capacity subsidy is excluded from consideration, i.e., at any budget, either a variable-capacity subsidy or a volume guarantee will result in higher incentivized capacity than a total-capacity subsidy.

Finally, we reinforce that a concessional loan instrument is equivalent to a total-capacity subsidy when the manufacturer savings and social investor costs of the loan instrument are equal. As illustrated with products B and C in Section 5, this is not always the case; a concessional loan becomes more attractive than a total-capacity subsidy when the loan cost to the social investor is less than the loan savings to the manufacturer.

Our study points to several future research opportunities. We note that a sales subsidy is less susceptible to moral hazard in settings where the manufacturer can meaningfully affect demand through its costly

[12] The total-capacity subsidy yields higher incentivized capacity for any budget than a sales subsidy, and neither the variable-capacity subsidy nor the volume guarantee will incentivize capacity investment at any budget.

actions. The manufacturer receives a payment for each unit sold under a sales subsidy. For the other instruments, the payment to the manufacturer is unaffected by sales. The issue of moral hazard is a worthy topic for future research. An additional worthy topic for future research is accounting for the possibility that the manufacturer has different beliefs regarding demand than the social investor. This issue, which can arise in practice, likely requires a significant and challenging new modeling dimension that seeks to characterize equilibria of a forecast signaling game.

While our study is most essential for manufacturing and distribution capacity needs for global health products in the LMIC markets, its application is broader as the key insights apply to developed markets. This can be seen from the COVID-19 pandemic: The question of how governments in developed nations can build manufacturing capacity in a rapid and effective manner is a public policy debate. Considering that the same set of concerns pose greater risks in LMIC markets, our study helps all social investors (e.g., government, development finance institution, philanthropic foundation) determine how best to utilize their financial resources for building capacity to treat diseases. Investments in such manufacturing capabilities can be perceived as an insurance in fighting pandemics and saving lives.

Acknowledgement. The authors are grateful to the editors and the reviewers whose comments greatly improved the paper. Burak Kazaz is grateful for the Whitman Roadmap Grant provided for this project at Syracuse University.

References

Aldy JE, Gerarden TD, Sweeney RL (2019) Investment versus output subsidies: implications of alternative incentives for wind energy. Working paper.

Alizamir S, de Véricourt F, Sun P (2016) Efficient feed-in-tariff policies for renewable energy technologies. *Operations Research* 64(1):52–66.

Arifoğlu K, Deo S, Iravani SMR (2012) Consumption externality and yield uncertainty in the influenza vaccine supply chain: Interventions in demand and supply sides. *Management Science* 58(6):1072–1091.

Atasu A, Van Wassenhove LN, Sarvary M (2009) Efficient take-back legislation. *Production and Operations Management* 18(3):243–258.

Atasu A, Subramanian R (2012) Extended producer responsibility for ewaste: individual or collective producer responsibility? *Production and Operations Management* 21(6):1042–1059.

Basu PK, Joglekar G, Rai S, Suresh P, Vernon JA (2008) Analysis of manufacturing costs in pharmaceutical companies. *Journal of Pharmaceutical Innovation* 3(1):30–40.

Barlow RE, Proschan F (1965) *Mathematical Theory of Reliability*. John Wiley & Sons, New York, NY.

Banciu M, Mirchandani P (2013) Technical note--new results concerning probability distributions with increasing generalized failure rates. *Operations Research* 61(4):925–931.

Bank D (2016a) Leveraging the balance sheet. *Stanford Social Innovation Research* 14(3):4–7

Bank D (2016b) Guaranteed impact. *Stanford Social Innovation Review* 14(3):16–18.

Berndt ER, Glennerster R, Kremer MR, Lee J, Levine R, Weizsacker G, Williams H (2007) Advance market commitments for vaccines against neglected diseases: Estimating costs and effectiveness. *Health Economics* 16(5): 491-511.

Chick SE, Mamani H, Simchi-Levi D (2008) Supply chain coordination and influenza vaccination. *Operations Research* 56(6):1493–1506.

Cohen MC, Lobel R, Perakis G (2016) The impact of demand uncertainty on consumer subsidies for green technology adoption. *Management Science* 62(5):1235–1258.

Dubois P, Lefouili Y, Straub S (2019) Pooled procurement of drugs in low and middle income countries. Toulouse School of Economics Working Paper TSE-999.

Erkoc M, Wu SD (2005) Managing high-tech capacity expansion via reservation contracts. *Production and Operations Management* 14(2):232-251.

Hill AM, Barber MJ, Gotham D (2018) Estimated costs of production and potential prices for the WHO Essential Medicines List. *BMJ Glob Health* 3(1):e000571.

Kania J, Kramer M, Russell P (2014 Summer). Strategic philanthropy for a complex world. *Stanford Social Innovation Review*.

Kazaz B, Webster S, Yadav P (2016) Interventions for an artemisinin-based malaria medicine supply chain. *Production and Operations Management* 25(9):1576–1600.

Kazaz B, Webster S, Yadav P (2021) Incentivizing COVID-19 vaccine developers to expand manufacturing capacity. *Center for Global Development Note*, March 26.

Kraiselburd S, Yadav P (2013) Supply chains and global health: an imperative for bringing operations management scholarship into action. *Production and Operations Management* 22(2):377–381.

Krass D, Ovchinnikov A, Nedorezov T (2013) Environmental taxes and the choice of green technology. *Production and Operations Management* 22(5):1035–1055.

Kremer M, Levin JD, Snyder CM (2022) Designing advance market commitments for new vaccines. *Management Science* to appear.

Lariviere M (2006) A note on probability distributions with increasing generalized failure rates. *Operations Research* 54(3):602–604.

Levine R, Pickett J, Sekhri N, Yadav P (2008) Demand forecasting for essential medical technologies. *American Journal of Law & Medicine* 34(2–3):225–255.

Martin P, Gupta D, Natarajan KV (2020) Vaccine procurement contracts for developing countries. *Production and Operations Management* 29(11):2601-2620.

Mitra S, Webster S (2008) Competition in remanufacturing and the effects of government subsidies. *International Journal of Production Economics* 111(2):287–298.

Natarajan K, Swaminathan JM (2014) Inventory management in humanitarian operations: impact of amount, schedule, and uncertainty in funding. *Manufacturing & Service Operations Management* 16(4):595–603

Özer Ö, Wei W (2006) Strategic commitments for an optimal capacity decision under asymmetric forecast information. *Management Science* 52(8):1238-1257.

Parish RM, McLaren KR (1982) Relative cost-effectiveness of input and output subsidies. *Australian Journal of Agricultural Economics* 26(1):1–13.

Park J, Kazaz B, Webster S (2018) Surface versus air shipment of humanitarian goods under demand uncertainty. *Production and Operations Management* 27(5):928-948.

Raz G, Ovchinnikov A (2015) Coordinating pricing and supply of public interest goods using rebates and subsidies. *IEEE Transactions on Engineering Management* 62(1):65–79.

Stiglitz JE (1987) Some theoretical aspects of agricultural policies. *The World Bank Research Observer* 2(1):43–60.

Song JS, van Houtum GJ, Van Mieghem JA (2020) Capacity and inventory management: review, trends, and projections. *Manufacturing & Service Operations Management* 22(1):36-46.

Tang CS, Wang Y, Zhao M (2015) The implications of utilizing market information and adopting agricultural advice for farmers in developing economies. *Production and Operations Management* 24(8):1197–1215.

Tantivess S, Chalkidou K, Tritasavit N, Teerawattananon Y (2017) Health technology assessment capacity development in low- and middle-income countries: experiences from the international units of HITAP and NICE. *F1000Research* 6:2119.

Taylor TA, Xiao W (2014) Subsidizing the distribution channel: donor funding to improve the availability of malaria drugs. *Management Science* 60(10):2461–2477.

ten Brink D, Gad M, Ruiz F (2018) Malaria innovations: pursuing value in an evolving market. *Lancet* 6(2):E138-E139.

Van Mieghem JA (2003) Capacity management, investment, and hedging: Review and recent developments. *Manufacturing & Service Operations Management* 5(4):269-302.

Webster S, Mitra S (2007) Competitive strategy in remanufacturing and the impact of take-back laws. *Journal of Operations Management* 25(6):1123–1140.

Webster S (2009) *Principles of Supply Chain Management*. Dynamic Ideas, Boston, MA.

Appendix (Online Supplement)

A. Regions of Instrument Cost Dominance

We assume that forecast error is normally distributed. Thus, uncertain demand can be expressed as

$$\tilde{d} = \mu + \tilde{z}\sigma$$

where $\tilde{z}$ is a standard normal random variable with pdf and cdf denoted as $\phi(z)$ and $\Phi(z)$. Therefore,

$$E\left[\tilde{d} \mid \tilde{d} \le x\right]F(x) = E\left[\mu + \sigma\tilde{z} \mid \tilde{z} \le \frac{x-\mu}{\sigma}\right]\Phi\left(\frac{x-\mu}{\sigma}\right) = \mu\Phi\left(\frac{x-\mu}{\sigma}\right) - \sigma\phi\left(\frac{x-\mu}{\sigma}\right)$$

$$E\min\left\{\tilde{d}, x\right\} = E\left[\tilde{d} \mid \tilde{d} \le x\right]F(x) + x\bar{F}(x) = \mu\Phi\left(\frac{x-\mu}{\sigma}\right) - \sigma\phi\left(\frac{x-\mu}{\sigma}\right) + x\left(1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\right)$$

(see Webster 2009). Let $m = p - c_v$ to simplify notation, $\alpha = F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$, $\alpha^o = F\left(x_2^o\right)$, and $\alpha_j^\tau = F\left(x_j^\tau\right)$. We replace x with α in our analysis (simplifies expressions), e.g., $j^*(\alpha) = \underset{j\in\{k,l,q\}}{\arg\min}\left\{b_j(\alpha)\right\}$ for all $\alpha \le \alpha_1^*$ in place of (18). (Note that if $x_2^* > 0$, then $\alpha_j^\tau = \min\left\{\alpha : \pi_j(\alpha) \ge 0\right\} = \alpha^o$ for $j \in \{s, k, l\}$.) Substituting into the profit expressions (see (1) and propositions 2 and 4), we obtain profit as a function of service level $\alpha > \alpha_j^\tau$,

$$\pi_s(\alpha) = \left(\frac{c_k}{1-\alpha}\right)\left(\mu\alpha - \sigma\phi\left(\Phi^{-1}(\alpha)\right)\right) - c_f$$

$$\pi_k(\alpha) = m\left(\mu\alpha - \sigma\phi\left(\Phi^{-1}(\alpha)\right)\right) - c_f \quad \text{if } m - c_f/\mu > 0$$

$$\pi_l(\alpha) = m\left(\mu\alpha - \sigma\phi\left(\Phi^{-1}(\alpha)\right) - (1-\alpha)c_f / c_k\right) \quad \text{if } m > 0$$

$$\pi_q(\alpha) = (m - c_k)\left(\mu + \sigma\Phi^{-1}(\alpha)\right) - c_f \quad \text{if } m - c_k > 0$$

$$\pi(\alpha) = m\left[\mu\alpha - \sigma\phi\left(\Phi^{-1}(\alpha)\right) + (1-\alpha)\left(\mu + \sigma\Phi^{-1}(\alpha)\right)\right] - c_k\left(\mu + \sigma\Phi^{-1}(\alpha)\right) - c_f .$$

Therefore, from (12), social investor cost as a function of service level $\alpha > \alpha_j^\tau = \min\left\{\alpha : \pi_j(\alpha) \ge 0\right\}$ can be expressed as

$$b_s(\alpha) = \pi_s(\alpha) - \pi(\alpha) = \left(\frac{c_k - m(1-\alpha)}{1-\alpha}\right)\left[\mu + (1-\alpha)\sigma\Phi^{-1}(\alpha) - \sigma\phi\left(\Phi^{-1}(\alpha)\right)\right]$$

$$b_k(\alpha) = \pi_k(\alpha) - \pi(\alpha) = \left(c_k - m(1-\alpha)\right)\left[\mu + \sigma\Phi^{-1}(\alpha)\right] \quad \text{if } m - c_f/\mu > 0$$

$$b_l(\alpha) = \pi_l(\alpha) - \pi(\alpha) = \left(c_k - m(1-\alpha)\right)\left[\mu + \sigma\Phi^{-1}(\alpha) + c_f / c_k\right] \quad \text{if } m > 0$$

$$b_q(\alpha) = \pi_q(\alpha) - \pi(\alpha) = m\sigma\left[\alpha\Phi^{-1}(\alpha) + \phi\left(\Phi^{-1}(\alpha)\right)\right] \text{ if } m - c_k > 0.$$

(Note that the functions cannot be analytically inverted to express α_j as function of b, e.g., no closed-form expression for $\Phi(z)$.) We include the expressions for instrument s above for completeness. However, the region where s dominates the other instruments is clear, i.e., s dominates if and only if $m \leq 0$ (which implies $x_2^* = 0$). The preferred instrument is less clear when $m > 0$. The following proposition identifies ability-to-pay indifference functions for each instrument pair among k, l, and q when $m > 0$.

Proposition A1. *Suppose that* $p - c_v > 0$. *Then*

(i) $p > c_v + c_f/\mu$: $b_k(\alpha) < b_l(\alpha)$ *iff* $\alpha > \alpha_k^\tau$

(ii) $p > c_v + c_k$: $b_q(\alpha) \leq b_l(\alpha)$ *iff at least one of the following conditions hold*:

$$\alpha > \alpha_q^\tau \text{ and } p \leq p_{ql}(\alpha) = c_v + \frac{c_k\left(\mu + \sigma\Phi^{-1}(\alpha)\right) + c_f}{\mu(1-\alpha) + \sigma\Phi^{-1}(\alpha) + \sigma\phi\left(\Phi^{-1}(\alpha)\right) + (1-\alpha)\frac{c_f}{c_k}}$$

$$\text{or } \alpha_l^\tau > \alpha > \alpha_q^\tau = \Phi\left(\frac{c_f/(p - c_v - c_k) - \mu}{\sigma}\right)$$

(iii) $p > c_v + \max\{c_f/\mu, c_k\}$: $b_q(\alpha) \leq b_k(\alpha)$ *iff at least one of the following conditions hold*:

$$\alpha > \alpha_q^\tau \text{ and } p \leq p_{qk}(\alpha) = c_v + \frac{c_k\left(\mu + \sigma\Phi^{-1}(\alpha)\right)}{\mu(1-\alpha) + \sigma\Phi^{-1}(\alpha) + \sigma\phi\left(\Phi^{-1}(\alpha)\right)}$$

$$\text{or } \alpha_k^\tau > \alpha > \alpha_q^\tau.$$

Proof. Part (i). If $x_2^* > 0$, then the result follows directly from Proposition 3; otherwise result follows from Proposition 4(ii) and (iv) and the fact that capacity functions $x_k(b)$ and $x_l(b)$ are strictly increasing in b, i.e., $x_k(b) = 0$ and $x_l(b) > 0$ for all $b \in \left[b_l^\tau, b_k^\tau\right]$ and $x_k(b) > x_l(b)$ for all $b > b_k^\tau > b_l^\tau$ imply $b_k(\alpha) <$ $b_l(\alpha)$ iff $\alpha > \alpha_k^\tau$.

Part (ii). Solving $b_q(\alpha) \leq b_l(\alpha)$ for p,

$$c_k\left[\mu + \sigma\Phi^{-1}(\alpha) + c_f/c_k\right] \geq m\sigma\left[\alpha\Phi^{-1}(\alpha) + \phi\left(\Phi^{-1}(\alpha)\right)\right] + m(1-\alpha)\left[\mu + \sigma\Phi^{-1}(\alpha) + c_f/c_k\right]$$

$$= m\left[\mu(1-\alpha) + \sigma\Phi^{-1}(\alpha) + \sigma\phi\left(\Phi^{-1}(\alpha)\right) + (1-\alpha)c_f/c_k\right] > 0$$

$$\Leftrightarrow \ p \le p_{ql}(\alpha) = c_v + \frac{c_k\left(\mu + \sigma\Phi^{-1}(\alpha)\right) + c_f}{\mu(1-\alpha) + \sigma\Phi^{-1}(\alpha) + \sigma\phi\left(\Phi^{-1}(\alpha)\right) + (1-\alpha)\frac{c_f}{c_k}} \ .$$

The expression for α_q^τ follows from the breakeven volume $x_q^\tau = c_f/(m - c_k)$.

Part (*iii*). Solving $b_q(\alpha) \le b_k(\alpha)$ for p,

$$c_k\left[\mu + \sigma\Phi^{-1}(\alpha)\right] \ge m\left[\alpha\sigma\Phi^{-1}(\alpha) + \sigma\phi\left(\Phi^{-1}(\alpha)\right)\right] + m(1-\alpha)\left[\mu + \sigma\Phi^{-1}(\alpha)\right]$$

$$= m\left[(1-\alpha)\mu + \sigma\Phi^{-1}(\alpha) + \sigma\phi\left(\Phi^{-1}(\alpha)\right)\right] > 0$$

$$\Leftrightarrow \ p \le p_{qk}(\alpha) = c_v + \frac{c_k\left(\mu + \sigma\Phi^{-1}(\alpha)\right)}{(1-\alpha)\mu + \sigma\Phi^{-1}(\alpha) + \sigma\phi\left(\Phi^{-1}(\alpha)\right)} \ . \ \square$$

To generate figures 4 and 5, we compute $\alpha^{o} = \dfrac{p - c_v - c_k}{p - c_v}$ and $\alpha_q^\tau = \Phi\left(\dfrac{c_f/(p - c_v - c_k) - \mu}{\sigma}\right)$ directly, and we numerically solve $\alpha_j^\tau = \min\left\{\alpha : \pi_j(\alpha) \ge 0\right\}$ for $j \in \{s, k, l\}$ to obtain α_j^τ. We numerically invert the indifference functions $p_{ql}(\alpha)$ and $p_{ql}(\alpha)$, for which the relevant portions of these curves appear in the figures.

B. Proofs

Proof of Proposition 2. Part (*i*). Observe that manufacturer profit $\pi_j(x, y)$ exhibits a newsvendor structure with respect to decision x. This observation underlies the following expressions for the manufacturer's capacity decision given that the participation constraint is ignored:

$$\hat{x}_s(y) = \arg\max_{x \ge 0}\left\{\pi_s(x, y)\right\} = \bar{F}^{-1}\left(\min\left\{\frac{c_k}{(p - c_v + y)^+}, 1\right\}\right)$$

$$\hat{x}_k(y) = \arg\max_{x \ge 0}\left\{\pi_k(x, y)\right\} = \bar{F}^{-1}\left(\min\left\{\frac{(1-y)c_k}{(p - c_v)^+}, 1\right\}\right),$$

which, given (6), reduce to (7) and (8).

Part (*ii*). We invert $\hat{x}_j(y)$ and substitute into $\pi_j(x, y)$ for $x > x_2^{o}$:

$$\hat{x}_s(y) = \bar{F}^{-1}\left(\frac{c_k}{p - c_v + y}\right) \qquad \Leftrightarrow \qquad y_s(x) = \frac{c_k}{\bar{F}(x)} - (p - c_v)$$

$$\pi_s(x) = \pi_s\left(x, y_s(x)\right) = \left(p - c_v + y_s(x)\right) E\min\left\{\tilde{d}, x\right\} - c_k x - c_f$$

$$=\left(\frac{c_k}{\bar{F}(x)}\right)E\min\left\{\tilde{d},x\right\}-c_kx-c_f=\left(\frac{c_k}{\bar{F}(x)}\right)\left(E\left[\tilde{d}\mid\tilde{d}\le x\right]F(x)+x\bar{F}(x)\right)-c_kx-c_f$$

$$=\left(\frac{c_k}{\bar{F}(x)}\right)E\left[\tilde{d}\mid\tilde{d}\le x\right]F(x)-c_f$$

$$\hat{x}_k(y)=\bar{F}^{-1}\left(\frac{(1-y)c_k}{p-c_v}\right)\qquad\Leftrightarrow\qquad y_k(x)=1-\left(\frac{p-c_v}{c_k}\right)\bar{F}(x)$$

$$\pi_k\left(x,y_k(x)\right)=(p-c_v)E\min\left\{\tilde{d},x\right\}-\left(1-y_k(x)\right)c_kx-c_f$$

$$=(p-c_v)E\min\left\{\tilde{d},x\right\}-\left(\frac{p-c_v}{c_k}\right)\bar{F}(x)c_kx-c_f$$

$$=(p-c_v)\left(E\left[\tilde{d}\mid\tilde{d}\le x\right]F(x)+x\bar{F}(x)-x\bar{F}(x)\right)-c_f$$

$$=(p-c_v)E\left[\tilde{d}\mid\tilde{d}\le x\right]F(x)-c_f\,.\;\square$$

Proof of Proposition 3. Part (*i*). From

$$\pi_s(x)=\left(\frac{c_k}{\bar{F}(x)}\right)E\left[\tilde{d}\mid\tilde{d}\le x\right]F(x)-c_f$$

(see Proposition 2), it is clear that $\pi_s(x)\to\infty$ as $x\to\infty$, which implies that x_s^τ is finite. Consequently, b_s^τ is finite. Furthermore,

$$\pi\left(x_2^\tau\right)=(p-c_v)E\left[\tilde{d}\mid\tilde{d}\le x_s^\tau\right]F\left(x_s^\tau\right)-c_kx_s^\tau-c_f<\mu\left(p-c_v-c_f/\mu\right)-c_kx_s^\tau<0$$

(due to $p-c_v-c_f/\mu\le 0$), and thus,

$$\infty>b_s^\tau=\pi_s\left(x_s^\tau\right)-\pi\left(x_s^\tau\right)=-\pi\left(x_s^\tau\right)>0.$$

From Proposition 2, for any $x>x_2^{o}$,

$$\pi_k(x)=(p-c_v)E\left[\tilde{d}\mid\tilde{d}\le x\right]F(x)-c_f\le\mu\left(p-c_v-c_f/\mu\right)\le 0$$

which implies $x_k(b)=0$ for all $b>0$, and $b_k^\tau=\infty$.

Part (*ii*). Following the arguments in the proof of Part (*i*), it follows that $0<b_s^\tau<\infty$ and $x_s^\tau>x_2^{o}$. Thus,

$$\pi'(x)<0 \text{ for all } x>x_2^{o}\tag{A1}$$

$$\bar{F}\left(x_s^\tau\right)<\bar{F}\left(x_2^{o}\right)=\frac{c_k}{p-c_v}$$

$$\pi_k\left(x_s^\tau\right)=\left(p-c_v\right)E\left[\tilde{d}\mid\tilde{d}\le x_s^\tau\right]F\left(x_s^\tau\right)-c_f=\frac{\bar{F}\left(x_s^\tau\right)}{c_k/\left(p-c_v\right)}\left(\frac{c_k}{\bar{F}\left(x_s^\tau\right)}\right)E\left[\tilde{d}\mid\tilde{d}\le x_s^\tau\right]F\left(x_s^\tau\right)-c_f$$

$$<\frac{\bar{F}\left(x_2^o\right)}{c_k/\left(p-c_v\right)}\left(\frac{c_k}{\bar{F}\left(x_s^\tau\right)}\right)E\left[\tilde{d}\mid\tilde{d}\le x_s^\tau\right]F\left(x_s^\tau\right)-c_f=\left(\frac{c_k}{\bar{F}\left(x_s^\tau\right)}\right)E\left[\tilde{d}\mid\tilde{d}\le x_s^\tau\right]F\left(x_s^\tau\right)-c_f$$

$$=\pi_s\left(x_s^\tau\right)=0.$$

Therefore,

$$b_s^\tau=\pi_s\left(x_s^\tau\right)-\pi\left(x_s^\tau\right)=-\pi\left(x_s^\tau\right)<\pi_k\left(x_s^\tau\right)-\pi\left(x_s^\tau\right)<\pi_k\left(x_k^\tau\right)-\pi\left(x_k^\tau\right)=-\pi\left(x_k^\tau\right)=b_k^\tau$$

and it follows that $0<b_s^\tau<b_k^\tau<\infty$, which in turn implies $0=x_k\left(b\right)<x_s\left(b\right)$ for all $b\in\left(b_s^\tau,b_k^\tau\right]$.

As noted in Section 4.1, the cost to incentivize investment in capacity x under instrument j (when viable) is the difference in manufacturer profits with, and without, the instrument (see (12)). Furthermore, for $x>x_2^o$,

$$\pi_k\left(x\right)=\frac{\bar{F}\left(x\right)}{c_k/\left(p-c_v\right)}\left(\frac{c_k}{\bar{F}\left(x\right)}\right)E\left[\tilde{d}\mid\tilde{d}\le x\right]F\left(x\right)-c_f<\left(\frac{c_k}{\bar{F}\left(x\right)}\right)E\left[\tilde{d}\mid\tilde{d}\le x\right]F\left(x\right)-c_f<\pi_s\left(x\right)\quad\text{(A2)}$$

which, in conjunction with (A1) and (12), implies $b_k\left(x\right)<b_s\left(x\right)$ for all $x\ge x_k^\tau$, and thus

$$0<x_s\left(b\right)<x_k\left(b\right)\text{ for all }b>b_k^\tau.$$

Part (*iii*). If $\pi\left(x_2^o\right)>0$, then $x_2^*=x_2^o>0$ and $b_s^\tau=b_k^\tau=0$ (see (11)). Then, from (A2), it follows that

$$0<x_2^*<x_s\left(b\right)<x_k\left(b\right)\text{ for all }b>0.\ \square$$

Proof of Proposition 4. Part (*i*). From the newsvendor structure of the manufacturer's profit function, it follows that

$$\hat{x}_l\left(y\right)=\arg\max_{x\ge0}\left\{\pi_l\left(x,y\right)\right\}=\bar{F}^{-1}\left(\min\left\{\frac{\left(1-y\right)c_k}{\left(p-c_v\right)^+},1\right\}\right)$$

$$\hat{x}_q\left(y\right)=\arg\max_{x\ge0}\left\{\pi_q\left(x,y\right)\right\}=\bar{F}^{-1}_{\max\left\{\tilde{d},y\right\}}\left(\min\left\{\frac{c_k}{\left(p-c_v\right)^+},1\right\}\right),$$

which, given (6), reduce to the expressions in Proposition 4(*i*).

Part (*ii*). We invert $\hat{x}_j\left(y\right)$ and substitute into $\pi_j\left(x,y\right)$ for $x>x_2^o$:

$$\hat{x}_l\left(y\right)=\bar{F}^{-1}\left(\frac{\left(1-y\right)c_k}{p-c_v}\right)\qquad\Leftrightarrow\qquad y_l\left(x\right)=1-\left(\frac{p-c_v}{c_k}\right)\bar{F}\left(x\right)$$

$$\pi_i(x)=\pi_i(x,y_i(x))=(p-c_v)E\min\{\tilde d,x\}-(1-y_i(x))(c_kx+c_f)$$
$$=(p-c_v)E\min\{\tilde d,x\}-\left(\frac{p-c_v}{c_k}\right)\bar F(x)(c_kx+c_f)$$
$$=(p-c_v)E\left[\tilde d\mid\tilde d\le x\right]F(x)-\left(\frac{\bar F(x)}{c_k/(p-c_v)}\right)c_f.$$

For instrument *q*, note that

$$\bar F_{\max\{\tilde d,y\}}(x)=\begin{cases}1, & x<y\\ \bar F(x), & x\ge y\end{cases}.$$

Therefore,

$$\hat x_q(y)=\arg\max_{x\ge 0}\left\{(p-c_v)E\min\{\max\{\tilde d,y\},x\}-c_kx-c_f\right\}=\begin{cases}x_2^o, & y\le x_2^o\\ y, & y>x_2^o\end{cases},$$

which implies

$$y_q(x)=x$$
$$\pi_q(x)=(p-c_v)E\min\{\max\{\tilde d,x\},x\}-c_kx-c_f=(p-c_v-c_k)x-c_f$$

for all $x>x_2^o$. □

Proof of Proposition 5. Part (*i*). From propositions 2 and, instrument $j\in\{k,l,q\}$ cannot incentivize investment in capacity (i.e., $\pi_j(x,y)<0$ for any $y>0$). Therefore, $b_k^\tau=b_l^\tau=b_q^\tau=\infty$ and $0=x_k(b)=x_l(b)=x_q(b)$. Since $\pi(x_2^o)<0$, $x_s^\tau=\pi_s^{-1}(0)>x_2^o$, and it follows that

$$b_s^\tau=b_s(x_s^\tau)=\pi_s(x_s^\tau)-\pi(x_s^\tau)=-\pi(x_s^\tau)\in(0,\infty)$$
$$0<x_s(b)\text{ for all }b>b_s^\tau.$$

Part (*ii*). First, from Proposition 3(*ii*),

$$0<b_s^\tau<b_k^\tau=\infty\text{ and }0=x_k(b)<x_s(b)\text{ for all }b>b_s^\tau.$$

Second, from Proposition 4, it follows from $p-c_v-\max\{c_k, c_f/\mu\}\le 0$ that

$$b_q^\tau=\infty\text{ and }x_q(b)=0\text{ for all }b>0.$$

What remains is to show that $b_l^\tau=b_s^\tau$ and $x_s(b)<x_l(b)$. From propositions 2 and 4, for any $x>x_2^o$,

$$\pi_s(x)=\left(\frac{c_k}{\bar F(x)}\right)E\left[\tilde d\mid\tilde d\le x\right]F(x)-c_f$$

$$\pi_l(x) = (p - c_v) E\left[\tilde{d} \mid \tilde{d} \le x\right] F(x) - \left(\frac{\bar{F}(x)}{c_k / (p - c_v)}\right) c_f$$

$$= \left(\frac{\bar{F}(x)}{c_k / (p - c_v)}\right) \left[\left(\frac{c_k}{\bar{F}(x)}\right) E\left[\tilde{d} \mid \tilde{d} \le x\right] F(x) - c_f\right] = \left(\frac{\bar{F}(x)}{c_k / (p - c_v)}\right) \pi_s(x)$$

Therefore,

$$\pi_s\left(x_s^\tau\right) = 0 = \pi_l\left(x_s^\tau\right),$$

which implies $x_l^\tau = x_s^\tau$ and $b_l^\tau = b_s^\tau \in (0, \infty)$. Finally, from $\pi_l(x) < \pi_s(x)$ for $x > x_2^{\mathrm{o}}$ (see (17)), it follows that

$$x_l(b) > x_s(b) > 0 \text{ for all } b > b_l^\tau = b_s^\tau.$$

Part (*iii*). From the proof of Part (*ii*),

$$0 < b_s^\tau = b_l^\tau < b_q^\tau = \infty$$

$$0 = x_q(b) < x_s(b) < x_l(b) \text{ for all } b > b_s^\tau = b_l^\tau,$$

and from Proposition 3(*ii*),

$b_s^\tau < b_k^\tau < \infty$ and $x_s(b) < x_k(b)$ for all $b > b_k^\tau$. The above results yield Part (*iii*).

Part (*iv*). From the proof of Part (*ii*),

$$0 < b_s^\tau = b_l^\tau < b_k^\tau = \infty$$

$$0 = x_k(b) < x_s(b) < x_l(b) \text{ for all } b > b_s^\tau = b_l^\tau$$

What remains is to show $b_q^\tau < \infty$. From Proposition 4,

$$\pi_q(x) = (p - c_v - c_k) x - c_f.$$

Thus, x_q^τ satisfies

$$\pi_q\left(x_q^\tau\right) = (p - c_v - c_k) x_q^\tau - c_f \Rightarrow x_q^\tau = \frac{c_f}{p - c_v - c_k} < \infty,$$

i.e., x_q^τ is breakeven volume. Note that $-\infty < \pi\left(x_q^\tau\right) < \pi\left(x_2^{\mathrm{o}}\right) < 0$ and

$$b_q^\tau = b_q\left(x_q^\tau\right) = \pi_q\left(x_q^\tau\right) - \pi\left(x_q^\tau\right) = -\pi\left(x_q^\tau\right) = -\left((p - c_v) E \min\left\{\tilde{d}, x_q^\tau\right\} - c_k x_q^\tau - c_f\right) < \infty.$$

Part (*v*). The results follow from the proofs of parts (*iii*) and (*iv*), i.e., *s*, *l*, and *k* results for the case of $p - c_v - c_f/\mu > 0$, and *s*, *l*, and *q* results for the case of $p - c_v - c_k > 0$. □

Proof of Proposition 6. Since $x_2^* = x_2^o > 0$ (i.e., capacity investment occurs without an instrument), it follows that

$$0 < \pi\left(x_2^o\right) = \left(p - c_v\right) E \min\left\{\tilde{d}, x_2^o\right\} - c_k x_2^o - c_f \le \left(p - c_v\right)\mu - c_k x_2^o - c_f \,,$$

which implies $p - c_v > c_f/\mu$, as well as $0 = b_s^\tau = b_k^\tau = b_l^\tau = b_q^\tau$ (see (11)). Furthermore, if $p - c_v - c_k \le 0$, then $\pi(x) \le 0$ for all x, which from $\pi\left(x_2^o\right) > 0$ implies $p - c_v > c_k$. Therefore, $p - c_v > \max\{c_k, c_f/\mu\}$.

From Proposition 3, $0 < x_2^* < x_s\left(b\right) < x_k\left(b\right)$ for all $b > 0$. Note that for $y_q = x_2^o$ ($= x_2^*$),

$$\begin{aligned} b_q\left(x_2^o\right) &= \pi_q\left(x_2^o\right) - \pi\left(x_2^o\right) = \left(p - c_v - c_k\right) x_2^o - c_f - \left[\left(p - c_v - c_k\right) E\left[\tilde{d} \mid \tilde{d} \le x_2^o\right] - c_f\right] \\ &= \left(p - c_v - c_k\right) E\left[x_2^o - \tilde{d} \mid \tilde{d} \le x_2^o\right] = \left(p - c_v\right) E\left(x_2^o - \tilde{d}\right)^+ \\ &= \left(p - c_v\right) E\left(\bar{F}^{-1}\left(\frac{c_k}{p - c_v}\right) - \tilde{d}\right)^+ > 0, \end{aligned}$$

and for incentivized capacity $x > x_2^o$, which requires volume guarantee $y_q(x) = x$ (see Proposition 4), we have

$$b_q\left(x\right) = \pi_q\left(x\right) - \pi\left(x\right) = \left(p - c_v\right) E\left(x - \tilde{d}\right)^+ .$$

Therefore, incentivizing capacity $x > x_2^o$ requires a budget larger than $b_q\left(x_2^o\right)$, which implies

$$x_q\left(b\right) = \begin{cases} x_2^o, & b \in \left[0, \left(p - c_v\right) E\left(x_2^o - \tilde{d}\right)^+\right] \\ b_q^{-1}\left(b\right), & b > \left(p - c_v\right) E\left(x_2^o - \tilde{d}\right)^+ \end{cases} .$$

For the remaining instruments $j \in \{s, k, l\}$, note that for $x > x_2^o$,

$$\begin{aligned} b_s\left(x\right) &= \pi_s\left(x\right) - \pi\left(x\right) = y_s\left(x\right) E \min\left\{\tilde{d}, x\right\} = \left(\frac{c_k}{\bar{F}\left(x\right)} - \left(p - c_v\right)\right)\left(E\left[\tilde{d} \mid \tilde{d} \le x\right] F\left(x\right) + x\bar{F}\left(x\right)\right) \\ &= \left(\frac{c_k}{\bar{F}\left(x\right)}\right) E\left[\tilde{d} \mid \tilde{d} \le x\right] F\left(x\right) + c_k x - \left(p - c_v\right) E \min\left\{\tilde{d}, x\right\} \end{aligned}$$

$$\begin{aligned} b_k\left(x\right) &= \pi_k\left(x\right) - \pi\left(x\right) = y_k\left(x\right) c_k x = \left(1 - \left(\frac{p - c_v}{c_k}\right)\bar{F}\left(x\right)\right) c_k x \\ &= c_k x - \left(p - c_v\right) x \bar{F}\left(x\right) \end{aligned}$$

$$b_l\left(x\right) = \pi_l\left(x\right) - \pi\left(x\right) = y_l\left(x\right)\left(c_k x + c_f\right) = \left(1 - \left(\frac{p - c_v}{c_k}\right)\bar{F}\left(x\right)\right)\left(c_k x + c_f\right)$$

$$= c_k x + \left(1 - \frac{\bar{F}(x)}{c_k / (p - c_v)}\right) c_f - (p - c_v) x \bar{F}(x).$$

Furthermore, $b_j(x)$ is strictly increasing in x and $\pi_k(x) < \pi_l(x) < \pi_s(x)$ for all $x > x_2^o$ (see (17)). Therefore $b_j(x)$ is invertible and $b_k(x) < b_l(x) < b_s(x)$ for all $x > x_2^o = x_2^*$, which due to $x_j(b) = b_j^{-1}(b)$, implies

$$0 < x_2^* < x_q(b) < x_s(b) < x_s(b) < x_k(b) \text{ for all } 0 < b \leq (p - c_v) E\left(x_2^* - \tilde{d}\right)^+$$

$$0 < x_2^* < x_s(b) < x_l(b) < x_k(b) \text{ for all } b > 0. \square$$

C. Mapping Annual Rate to Investment Horizon Rate

The net present value (NPV) of annuity of amount $1/T$ that is paid at the end of year 1, 2, …, T at annual rate r (for a total payout of \$1) is

$$N_1 = \frac{1}{T}\left(\frac{(1+r)^T - 1}{r(1+r)^T}\right).$$

The net present value of a payout \$1 at time T_2 and annual rate r is

$$N_2 = \frac{1}{(1+r)^{T_2}} .$$

Setting $N_1 = N_2$ and rearranging,

$$i_2 = (1+r)^{T_2} - 1 = \left(\frac{rT}{(1+r)^T - 1}\right)(1+r)^T - 1 .$$

Thus, the NPV of total payoff P with equal payments at the end of each year for T years from investment I at time 0 is equal to the NPV of total payoff P received at time T_2 from investment I at time 0, i.e.,

$$N_2 P - I = \frac{P}{1+i_2} - I = \frac{P}{1 + \left((1+r)^{T_2} - 1\right)} - I = \frac{P}{T}\left(\frac{(1+r)^T - 1}{r(1+r)^T}\right) - I = N_1 P - I .$$