POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT

POMS

Vol. 29, No. 3, March 2020, pp. 571–594
ISSN 1059-1478 | EISSN 1937-5956 | 20 | 2903 | 0571

DOI 10.1111/poms.13129

Value of Combining Patient and Provider Incentives in Humanitarian Health Care Service Programs

Mili Mehrotra
University of Illinois at Urbana-Champaign, Champaign, Illinois 61820, USA, milim@illinois.edu

Karthik V. Natarajan*
University of Minnesota, Minneapolis, Minnesota 55455, USA, knataraj@umn.edu

The below-par progress toward the Millennium Development Goals in many developing countries has been attributed to the low availability of good quality health care services and to the demand-side barriers to access. In this study, we analyze an incentive design problem faced by a budget-constrained humanitarian organization managing a health care service program (e.g., maternal health or HIV services) with different emphasis on the two measures of quality—structural and process. Incentives offered to the health care provider (referred to as supply-side incentives) are aimed at improving the availability of good quality services and demand-side incentives are used to encourage patients to seek care. In many developing country health programs, incentive schemes tend to be purely supply-side focused or purely demand-side focused. However, our results suggest that by offering the right combination of incentives to the patients and the provider, program performance can be increased up to 20 times, on average, depending on the service offered. Even within the current practice of using one-sided incentives, there is significant scope for improvement (as much as 14 times, on average) by ensuring that there is better alignment between the incentive scheme and the service offered. In particular, pure supply-side incentive schemes perform better than pure demand-side incentive schemes in programs where there is a higher emphasis on structural quality. The opposite result holds true for programs that place a higher emphasis on process quality.

Key words: incentive design; health care services; humanitarian operations
History: Received: June 2018; Accepted: October 2019 by Michael Pinedo after two revisions.

1. Introduction

The last two decades have witnessed unprecedented attention to health challenges in developing countries. Development Assistance for Health, which is the aid funding given to support health-related goals in developing countries, has increased from \$5.7 billion in 1990 to about \$38.9 billion in 2018 (Chang et al. 2019). The manifold increase in funding has led to impressive gains in several health indicators. However, the gains have not been uniform across countries and health conditions, with several developing countries falling short of the Millennium Development Goals (MDGs) adopted by the UN Declaration in 2000 (The World Bank 2015).

In developing countries, supply-side constraints have frequently been identified as the limiting factors that hamper the progress toward health targets. These constraints include limited availability and low quality of health care services, weak infrastructure, and lack of trained workforce. Consequently, efforts to improve the performance of health programs in developing countries have traditionally focused on designing supply-side interventions that are aimed at improving the availability and quality of health care services offered (Ensor and Cooper 2004). However, the supply side is only one part of the equation. Demand-side barriers such as lack of awareness and the opportunity cost of seeking care also play a key role, especially in developing country settings, and interventions to improve the performance of health care programs can lead to underwhelming results if the demand-side barriers are not fully taken into account. For example, according to Amudhan et al. (2013), strengthening a primary health center network in India resulted in a 14.5% increase in institutional deliveries at villages close to the health center but the rate of increase was half that number in villages with poor access to this center. This example and several other works (see e.g., Mangham and Hanson 2010, Yamey 2012) highlight the critical role played by demand-side constraints in determining a patient's decision to seek or to not seek health care services in developing countries. To alleviate the negative impact of demand-side barriers, several interventions have been employed in practice including conducting awareness campaigns, providing transportation vouchers, and conditional cash transfers to reward patients who adhere to treatment protocols (Ensor and Cooper 2004).

Similar to the shortcomings of pure supply-side strategies that we just highlighted, there are inherent limitations associated with using pure demand-side strategies as well. Clearly, generating additional demand will not translate into better health outcomes unless the health program has adequate resources to meet the increased demand for services offered by the program. For instance, in Sierra Leone, a local city government announced the extension of free health care services to all pregnant and breastfeeding women in 2010. However, the city's health system lacked the capacity and resources to meet the increased demand, resulting in frustration and an unfulfilled promise to offer free maternal care (Fofanah 2010).

Based on this discussion, it is clear that investing resources to increase the availability and quality of health care services and creating incentives to encourage people to seek care are both vital to achieving improvements in health outputs and outcomes. However, in developing countries, public health agencies and not-for-profit organizations play a major role in managing health care programs and they typically operate under a limited budget. Given the limited funding availability, a key decision is to determine how much funding should be allocated toward improving the availability and quality of service offered and how much funding should be allocated toward offering incentives to patients to encourage them to seek care.

In this study, we consider a setting where a budget-constrained humanitarian organization[1] manages a health care program that offers a particular service (e.g., maternal and child health services, family-planning, HIV services) to a target population. The organization's objective is to maximize the *quality-adjusted program coverage*, subject to the limited funding availability. Here coverage refers to the number of people served by the program. In a health care context, there are several ways to define quality including the pioneering framework presented by Donabedian (2005). According to the Donabedian model, there are three key components of health care quality: structure, process, and outcome. "Structure" refers to the physical and infrastructure-related aspects of providing care and examples of structural quality measures include the availability of equipment, medical supplies, technology, human resources, and management capabilities (in terms of having policies and procedures in place to ensure consistency of care). Broadly speaking, the structural elements create a more conducive environment for providing good medical care but depending on the service offered, they may not necessarily guarantee better health outcomes.

"Process" measures capture the quality of interaction between patients and a health care provider, and determine the clinical content of care (e.g., how the care was delivered including its appropriateness and completeness, compliance with recommended clinical guidelines, and providing patients with information and encouraging them to be a part of the decision-making process). For instance, in case of prenatal care, process measures could include examination of medical history, discussing pregnancy-related risks, and planning the delivery in consultation with the patient. Provided all the tangible aspects remain the same, it is reasonable to expect that the clinical content of care will improve with the provider's willingness to spend more time with each patient. Finally, "outcome" refers to the end result of the care process and is captured by measures such as readmission rates, mortality etc. Clearly, *both structure and process measures play an important role in determining the outcomes* but as highlighted by Donabedian (2005) and several others (Boyce 1996, Hayford and Maeda 2017), there are challenges associated with developing and using outcome measures. For example, many factors other than medical care (that are outside the control of the provider) may influence outcomes and in some cases, it may take months or even years before the relevant outcomes are observed (Donabedian 2005). Monitoring and controlling for external factors that could influence outcomes is likely to be particularly challenging in environments that are typical of many humanitarian health care programs. As a result, quality evaluations in most programs focus on structure and process measures (see e.g., Gertler and Vermeersch 2012, USAID 2016b,c) and consistent with those programs, we focus on those two measures of quality as well.

In several developing countries, service delivery is often carried out by private health care providers and the focus of our study is on such settings. In recent years, humanitarian organizations have increasingly turned toward performance-based contracts for health care providers instead of paying them based on inputs (The AIDSTAR-2011, 2011, World Bank Group 2015). Consistent with practice, we assume that the humanitarian organization evaluates the provider on both structural and process quality measures for performance-based payments (Gertler and Vermeersch 2012, USAID 2016b,c). The humanitarian organization assigns weights to structural and process quality, which capture the organization's perceived importance of structural and process measures as it relates to the quality of care offered by the provider. The overall quality of care is calculated as a weighted sum of the structural and process quality scores. In this study, we focus attention on linear performance-based payment schemes where the service provider is paid a per-unit amount, adjusted by the quality of care offered, for every patient treated. Such aggregate-quality based, linear payment schemes have been used extensively in humanitarian health care programs including maternal and child health programs in Rwanda and Nigeria (see Gertler and Vermeersch 2012 and USAID 2016a respectively) primary care in Senegal (USAID 2016b) and for infection prevention and HIV services in Mozambique (USAID 2016c). In addition to being practically relevant, our extensive numerical experiments indicate that the aggregate-quality based payment scheme also performs well relative to a benchmark centralized model where the humanitarian organization directly controls the provider (performance gap is less than 2% in nearly 81% of the instances).

On the demand side, we consider monetary incentives to encourage patients to seek care. More specifically, we focus on per-capita incentives offered to people within the catchment area to partially or fully offset the cost of seeking treatment. Per-capita incentive schemes are widely used in humanitarian health care programs and they are implemented in several ways including vouchers for family planning and maternal health in more than a dozen countries (Grainger et al. 2014) and in the form of conditional cash transfers for vaccination and maternal health in various countries including Bangladesh, Pakistan, India and Mexico (Jehan et al. 2012, Lagarde et al. 2007).

We capture the strategic interaction between the humanitarian organization, the health care provider, and patients using a three-stage sequential decision-making model. In practice, depending on the health care service offered by a program (lab testing vs. family planning, for example), the humanitarian organization may place varying emphases on structural and process quality measures. Hence, a key focus of our research is to understand how the type of health care service (structure-focused or process-focused) influences the incentives offered to patients and the health care provider. In particular, we address the following issues of interest to budget–constrained organizations managing health care programs in developing countries:

1. *Analyzing incentive schemes used in practice:* How do the one-sided incentive schemes that focus exclusively on either the supply side or the demand side perform relative to one another for different types of health care services? How do the incentives offered in such one-sided incentive schemes change based on the available budget, type of health care service and the program's operating environment (i.e., patient and provider characteristics)?
2. *Two–sided incentive design:* Given a limited budget, what is the right combination of incentives to the patients and the provider in order to maximize quality-adjusted coverage? How do the incentives change based on the available budget, type of health care service and the program's operating environment?
3. *Value of combining patient and provider incentives:* Under what situations does offering the right combination of patient and provider incentives lead to significant benefits over one–sided incentive schemes used in practice, and under what conditions would offering incentives exclusively to either the provider or patients suffice?

To address these questions, we analyze our model theoretically as well as numerically using a realistic test bed. We next describe the key insights from our analysis.

With respect to the one-sided incentive schemes used in practice, our key finding is that humanitarian organizations would benefit from using pure supply-side incentive schemes over demand-side focused incentive schemes in health care programs where structural quality is more important. The opposite result holds true for programs where process quality might be deemed more important and in fact, we demonstrate that under certain conditions, using a pure demand-side incentive scheme is optimal for exclusively process-focused programs. An important implication of these results is that even within the current practice of using one-sided incentives, there is significant scope for improving the performance of humanitarian health care programs (as much as 14 times) by ensuring that there is better alignment between the incentive scheme and the type of service offered.

Our theoretical analysis also brings to light similarities as well as important differences between one-sided and two-sided incentive schemes. We find that the patient incentive always increases with available funding under pure demand-side incentive schemes and in a similar vein, the provider incentive also increases, at least weakly, with available funding under pure supply-side incentives. When using a two-sided incentive scheme, the incentive offered to patients increases with available funding but contrary to conventional wisdom (and in contrast to the result under pure supply-side incentive schemes), it may not be always optimal to increase the incentive offered to both patients and the provider simultaneously as additional funds become available. We demonstrate that in certain situations, an increase in patient incentive should be accompanied with a reduction in the incentive offered to the provider to achieve maximal impact. With respect to the health care service offered, we find that under pure supply-side incentive schemes, the humanitarian organization should lower the provider incentive as the emphasis on structural quality goes up. The impact of process quality is, however, more nuanced. We find that, as the emphasis on process quality goes up, it is optimal for the humanitarian organization to lower the provider incentive in settings where structural quality improvements lead to a large increase in the provider's capacity. However, this result may not necessarily hold true in settings where the impact of structural quality investments on the provider's capacity is relatively low. A similar effect also holds true under two-sided incentives in terms of how the provider's structural quality investments change in response to an increasing emphasis on process quality. These results highlight the need for humanitarian organizations to exercise caution and understand the underlying operating environment before transferring learnings related to effective incentive design from one health care program to another.

Our extensive numerical study brings forth the following insights regarding the value of optimal incentive design that could be of interest to humanitarian health care organizations operating in developing countries. We find that by offering the right combination of incentives to the patients and the provider (as opposed to using one-sided incentive schemes), humanitarian organizations can achieve a significant improvement in program performance (up to 20 times, on average, depending on the budget and type of service offered). Using the right combination of incentives is also critical to fully realizing the benefits of additional funding—fundraising efforts will have little to no impact on program performance unless the appropriate incentives are in place for patients to seek care and for providers to offer the best possible care. Finally, our analysis also reveals that by switching from one-sided incentive schemes to two-sided incentives, humanitarian organizations can provide better quality treatment to more people despite having less funding. In recent years, developing country health programs have been confronted with the dilemma of a growing demand for health commodities and services at a time when the aid funding received from developed nations is beginning to either stagnate or even decline in some cases (Leach-Kemon et al. 2012). Our results suggest that the improved aid effectiveness resulting from using the right incentive scheme could play a critical role in helping humanitarian organizations to successfully address this challenge.

2. Literature Review

Our work contributes to the growing literature on managing socially responsible and humanitarian operations (Alizamir et al. 2019, Chen et al. 2019, Gallien et al. 2017, Mu et al. 2016, Natarajan and Swaminathan 2014, Privett and Erhun 2011). Within this stream, the existing papers have studied a variety of issues but the ones that are most relevant to our work are resource allocation and incentive design. Jonasson et al. (2017) focus on capacity allocation decisions that arise in the context of infant HIV diagnosis. They develop a two-stage model to analyze how the assignment of clinics to labs and the allocation of available diagnostic capacity across labs impacts the test sample turnaround times and the number of infants initiating treatment. Arora et al. (2017) consider a funding allocation decision that arises in not-for-profit service settings. They analyze how a not-for-profit organization with limited financial resources should allocate the available funding between advisory efforts (to direct clients to services that are most appropriate for their needs) and service delivery efforts. Yang et al. (2013) develop an optimization model to decide which children, from among a group, should receive ready-to-use therapeutic food (RUTF) or supplementary food. In their model, the decision maker has limited funding availability that can be used to procure RUTF and/or supplementary food.

Paul et al. (2019) analyze contracting and incentive design challenges that arise in blood bank operations. Due to the unpredictable and limited supply of blood products, blood banks are often forced to ration the available supply between different hospitals, which in turn encourages them to inflate their orders. This leads to a loss of welfare across the entire blood supply chain. The authors demonstrate that by offering a subsidy for every unit of shortage experienced by the hospitals, blood banks can induce them to order their true demand without inflation. Berenguer et al. (2017) and Taylor and Xiao (2014) analyze the impact of sales and purchase subsidies on the consumption/sales of a socially beneficial product (for example, malaria bed nets). In both these papers, the decision maker has limited funding available that can be utilized to provide subsidies. Motivated by the influenza vaccine industry, Dai et al. (2016) study a contracting problem between a manufacturer and retailer to improve the on-time delivery of the vaccines. Due to uncertainties in the product composition and demand, manufacturers are often reluctant to engage in early production, resulting in vaccine shortages that have significant negative societal implications. Dai et al. (2016) perform a comparative analysis of the different contracts currently used in practice and propose a new form of contract that has the potential to improve the on-time delivery of influenza vaccines.

The papers mentioned above focus on resource allocation and incentive design from *a supply-side perspective*. In contrast, our paper looks at the problem of using available funding to offer both supply- and demand-side incentives in a humanitarian health care service setting. In terms of the emphasis on allocating resources to influence both supply and demand in humanitarian health care settings, a relevant work is Natarajan and Swaminathan (2019). However, they analyze settings involving health products while our focus is on health care services. In Natarajan and Swaminathan (2019), the humanitarian organization manages the procurement of the health product while demand mobilization is contracted to local agents. In our work, demand mobilization does not involve contracting and is done directly through monetary incentives to patients. In addition, we also consider performance-based incentives on the supply side, a feature that is not present in their paper.

3. Model

We consider a budget-constrained humanitarian organization managing a health care program that offers a particular service to patients within a community. Examples of services offered by humanitarian health care programs in developing countries include maternal and child health services, family planning, and HIV services to name a few. The organization's objective is *to identify the right level of incentives to offer to the patients and the health care provider* so as to maximize the "quality-adjusted" program coverage, given the limited funding availability.

We model the strategic interaction between the humanitarian organization, health care provider, and patients using a three-stage sequential decision-making framework. In the first stage, the humanitarian organization decides the incentive amount to be offered to patients (denoted by P per patient) and the health care provider (denoted by p per patient). Considering these incentives as given and anticipating the patients' response, the provider makes decisions regarding service speed and the quality of care offered to patients. Finally, patients decide whether or not to seek care, depending on the incentive offered to them by the humanitarian organization and the provider's decisions. This sequence of events is also illustrated in Figure 1. In what follows, we describe the details of the three-stage sequential process.

3.1. Patient's Decision: To Seek or To Not Seek Care?

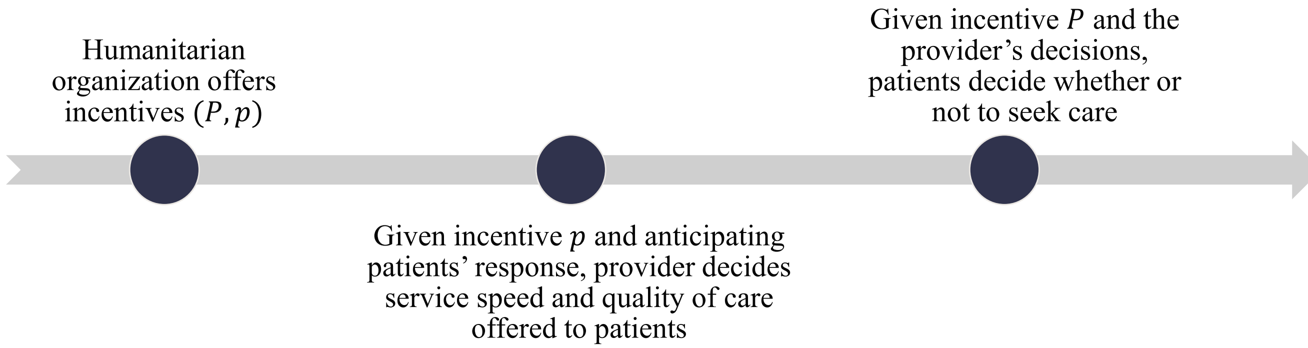
As we discussed in the introduction, in developing countries, demand-side barriers significantly impact a patient's decision to seek or to not seek care, even when the service is offered free of cost. Important demand-side barriers include a lack of awareness of the availability and benefits of seeking care, transportation costs to get to a health center, and the opportunity cost of the time spent to receive care. In our model, V_b represents a patient's perceived value of getting treatment, which is closely tied to the population awareness level. For simplicity, we assume that the fixed cost of transportation to the health center is the same for all patients (denoted by C_t). When deciding whether or not to seek care, patients also consider the expected opportunity cost, C_w, associated with waiting and receiving treatment at the health center.

Several factors impact a patient's expected waiting time including the provider's service rate and the arrival rate of patients to the health center. Following the modeling framework used by several papers within the operations management literature (Anand et al. 2011, Kostami and Rajagopalan 2014), we consider an M/M/1 queuing system to model the interplay between the provider's service rate and the patients' decision to seek or to not seek care. For a M/M/1 model, the patient's expected opportunity cost is given by $C_w(\mu, \lambda_e) = c_w W(\mu, \lambda_e)$, where $W(\mu, \lambda_e) = \frac{1}{(\mu - \lambda_e)}$. Here $c_w > 0$ represents the patient's opportunity cost per unit time, μ is the provider's mean service rate, and λ_e is the effective arrival rate of patients to the health center, a key measure of interest to the humanitarian organization.

The organization can increase the number of people seeking access to the health care service through demand-side interventions. In our work, we focus on per-capita incentives offered to patients, a form of demand-side intervention that is frequently used in humanitarian health care programs in developing countries. The per-capita incentive helps to offset some of the costs and disutility associated with seeking care. Given per-capita incentive P, a patient will seek care if the perceived net utility from obtaining treatment is positive, that is, if $V_b + P - C_t - c_w W(\mu, \lambda_e) \geq 0$. The effective arrival rate of patients to the health center is then given by $\lambda_e = \mu - \frac{c_w}{V_b + P - C_t}$. To maintain stability of the M/M/1 system, we assume that $V_b - C_t > 0$. When combined with $c_w > 0$, this ensures that $\mu - \lambda_e > 0$ for all finite values of P. Notice that the arrival rate expression does not include any payments made by patients, that is, we assume that the program offers the service free of cost. Having a fixed service fee would not structurally alter our results.

In addition to the per-capita incentive, the patient arrival rate could also be influenced by the quality of care offered by the provider. Specifically, an increase in either structural or process quality or both could increase the patient's perceived value of getting treatment (i.e., V_b). This in turn could result in higher arrival rates. Furthermore, as we discuss in the next section, structural quality improvements also have the potential to indirectly increase the patient arrival

Figure 1 Sequence of Events [Color figure can be viewed at wileyonlinelibrary.com]



rates through their impact on the provider's capacity. Specifically, in some cases, investments in structural quality lead to an increase in the provider's capacity, which in turn could increase the patient arrival rate due to the reduction in waiting times. We incorporate this influence of structural quality on the provider's service capacity in our model, as outlined in the next section. However, explicitly including the effect of structural and process quality into the patient valuation V_b makes the model intractable and leads to significant analytical challenges. Hence, we do not include them in our main model. Nevertheless, we conducted an extensive numerical study to explore the effect of including structural and process quality into the patient's valuation and our results indicate that all the key insights obtained from our main model (pertaining to quality-adjusted coverage and the patient and provider incentives) continue to hold in this more complex setting as well. We describe the modified model under "model extension" in section 7.1.

3.2. The Provider's Optimal Response

Now we focus on the decisions made by the health care provider, who is a *profit maximizer*. As mentioned earlier, there are three dimensions of health care quality, namely structure, process and outcome. However, due to the challenges in measuring and using outcome measures, humanitarian organizations typically evaluate the provider based on structural and process quality measures. We use $R_s \in [0, 1]$ and $R_p \in [0, 1]$ to denote the provider's structural and process quality scores, respectively. We wish to clarify that $R_s = 0$ does not imply that the provider has no physical and infrastructural resources required to offer care. Rather, it corresponds to a scenario where the provider does not invest in resources that would structurally enhance the environment in which care is offered beyond what is deemed "minimally essential" by the humanitarian organization. Similarly, $R_p = 0$ implies that the clinical content of care is at a "minimally acceptable" level as per pre-established treatment guidelines, and the provider does not engage in activities that would be viewed by the humanitarian organization as enhancing the patient's clinical experience. The humanitarian organization assigns weights w_s and w_p to structural and process quality scores, respectively. The provider's overall quality score is given by $R = w_sR_s + w_pR_p$.

In our work, we assume that the organization offers a "quality-adjusted" payment contract where the provider receives payment p per patient, multiplied by the provider's aggregate quality score $R \in [0,1]$, to yield a total payment of pR for every patient seen. Making provider payments contingent on the aggregate score R is consistent with practice and this approach is used across a variety of humanitarian health care programs in different countries including maternal and child health services in Rwanda and Nigeria, primary care in Senegal, and infection prevention, maternal care and HIV services in Mozambique (see Gertler and Vermeersch 2012, USAID 2016a,b, and USAID 2016c for detailed descriptions of the payment schemes used in Rwanda, Nigeria, Senegal, and Mozambique, respectively). A common theme across the above-mentioned humanitarian programs is that the quantity-based payments (=$p\times$ number of patients seen) are scaled by the aggregate quality score of the provider. Given its widespread use in practice, we also consider aggregate-quality-based provider payments in our paper. We return to this issue and provide further justification for the use of aggregate-quality based payments in section 3.3.

The humanitarian organization sets the per-unit payment p but the provider can control the actual payment received per patient (=pR) by influencing the quality of service offered. More specifically, the provider can improve the structural quality score R_s through appropriate investments I in infrastructure and facility-related characteristics. In practice, structural quality evaluations place a heavy emphasis on the availability of medical equipment, drugs and qualified health personnel necessary to deliver adequate care. As such, an increase in the availability of any of the above-mentioned resources could increase the provider's capacity to offer service. We incorporate this into our model by assuming that the provider's service rate increases by γR_s with structural quality investments, where $\gamma \geq 0$ represents the sensitivity of the provider's capacity to those investments. The higher capacity, in turn, benefits patients by lowering their wait times and/or enabling the provider to offer better process quality to patients (this aspect will become clear in the ensuing discussion on process quality). Based on the literature on quality improvement in health care settings (see Gandjour 2010), we make the intuitive assumption that structural quality investments exhibit a diminishing rate of return. We model the diminishing returns using the function $I = \beta R_s^2$ where $\beta > 0$ is the scaling factor. Quadratic functions have been used by several papers to capture increasing, convex behavior (e.g., Heese and Swaminathan 2006).

In addition to structural quality, the provider can also influence the process quality score R_p by appropriately adjusting the service time. To understand the relationship between service rate and process quality, we need to first look at the impact of changing the service rate. Increasing the service rate could have a positive effect on the waiting time and the number of people seen by the provider but in customer-intensive services like health care, it has been well documented in the operations management literature (Anand et al. 2011, Kostami and Rajagopalan 2014, Wang et al. 2019) as well as in global health settings (Silverman et al. 2015) that the quality of service inevitably decreases with service speed. This leads to a quality-speed conundrum for the provider since in many health care settings including ours, the payments to the provider depend directly on both the *quality of service offered and the number of people served*. In such cases, the provider chooses a service rate to optimally balance the quality-speed trade-off but clearly, the provider's actions also impact the patient's waiting time and the opportunity cost of seeking care and consequently, the patient arrival rate.

Similar to the linear framework used in the literature (see Anand et al. 2011, Kostami and Rajagopalan 2014), we assume that the provider's mean service rate is $\mu = \alpha + \gamma R_s - aR_p$, that is, the service rate goes down as the provider seeks to improve process quality. Notice that when $R_s = R_p = 0$, the provider's service rate is α. Hence, we refer to α as the provider's base capacity (expressed in patients served per unit time). Parameter a ($0 \leq a \leq \alpha$) captures the sensitivity of service speed to process quality. Notice that $\mu = \alpha + \gamma R_s - aR_p$ implies that the mean service time τ equals $\frac{1}{\alpha+\gamma R_s - aR_p}$, that is, process quality is increasing and concave in the service time. The expression for μ also highlights the two key benefits of structural quality investments that we discussed earlier: (i) the additional capacity (γR_s) could be used to increase the service rate (and reduce patient wait times) while maintaining process quality at its current level or (ii) it could be used to offer better process quality to patients without a reduction in the number of patients seen. Both of these in turn influence the quantity and quality of care offered by the humanitarian health care program.

In our analysis, for convenience, we let R_s and R_p be the provider's decisions. Before presenting the provider's optimal responses[2] R_s^* and R_p^*, we describe the provider's decision-making sequence. Structural aspects of quality typically depend on long-term decisions (e.g., purchasing new equipment, hiring new personnel or developing better program management capabilities) that require advance planning while service-time adjustments are more flexible and short-term in nature. Hence, we assume that the provider first makes investment decisions that impact structural quality. Then, taking the structural aspects of quality as fixed, the provider makes service-time adjustments to influence the process quality score. Table 1 summarizes our key notation. We introduce other notation as and when needed.

We now present the profit maximization problem that drives the provider's process and structural quality choices. For any given incentives p and P, the provider's problem is the following.

$$\max_{0 \leq R_p, R_s \leq 1} \pi = pR\lambda_e - I(R_s) = p\left(w_s R_s + w_p R_p\right) \left(\alpha + \gamma R_s - aR_p - \frac{c_w}{V_b + P - C_t}\right) - \beta R_s^2 \tag{1}$$

Table 1 Model Parameters and Decision Variables

Parameters	
Organizational characteristics	
B	Funding available to the humanitarian organization
Provider characteristics	
α	Provider's base capacity (in patients/unit time)
γ	Sensitivity of the provider's capacity to structural quality investments
Patient characteristics	
V_b	Patient's perceived value of getting treatment
C_t	Fixed cost of transportation to the health center
c_w	Patient's opportunity cost per unit time
Program characteristics	
w_s	Weight assigned to structural quality
w_p	Weight assigned to process quality
Decision variables	
Provider's decisions	
R_s	Structural quality score
R_p	Process quality score
Humanitarian organization's decisions	
p	Incentive offered to the provider
P	Per-capita incentive offered to patients

In expression (1), notice that the provider's revenues ($pR\lambda_e$) are per unit time so naturally the investment I to improve structural quality (and the associated scaling factor β) are also expressed on a per unit time basis – investments toward purchasing equipment and training-related costs are assumed to be amortized. Throughout our analysis, we assume that $a + \frac{c_w}{V_b - C_t} \leq \alpha < 2a$ and $\frac{aw_s}{w_p} \leq \gamma \leq 2a - \alpha$. Condition $a + \frac{c_w}{V_b - C_t} \leq \alpha$ ensures that demand (λ_e) is non-negative regardless of the provider's decisions. Conditions $\alpha < 2a$ and $\frac{aw_s}{w_p} \leq \gamma \leq 2a - \alpha$ allow us to concisely present our theoretical results. We relax the latter two assumptions in our computational study. In expression (1) the term $\alpha + \gamma R_s - aR_p - \frac{c_w}{V_b + P - C_t}$ is the patient arrival rate. Notice that the arrival rate is indirectly influenced by the structural quality investments through its impact on the provider's service rate. In addition to this effect, the arrival rate could also be positively influenced by improvements in structural and process quality since patients' perceived value of getting treatment, V_b, may increase with the quality of care offered by the provider. We consider this model extension in section 7.1.

Returning to expression (1), we use backward induction to characterize the provider's optimal response for any given incentives p and P. The results are summarized in Lemma 1. The proofs of all the results are provided in the Online Appendix. For brevity, we define $F(P) = \alpha - \frac{c_w}{V_b + P - C_t}$ and use this notation throughout the study. $F(0)$ simply refers to the function evaluated at $P = 0$.

Lemma 1. *For any given* (p,P), *the provider's structural and process quality choices are as follows:*

$$R_s^* = \min\left\{\frac{pw_p(a\frac{w_s}{w_p} + \gamma)F(P)}{4a\beta - pw_p(a\frac{w_s}{w_p} + \gamma)^2}, 1\right\},$$
$$R_p^* = \frac{1}{2a}\left(F(P) + \left(\gamma - \frac{aw_s}{w_p}\right)R_s^*\right).$$

Lemma 1 enables us to understand how changes in incentives impact the provider's quality choices and the arrival rate of patients to the health center. Proposition 1, derived from Lemma 1, summarizes the impact of changing the incentive offered to the provider while keeping the patient incentive fixed.

Proposition 1. *For any given patient incentive* P *and combination of system parameters,* R_s^*, R_p^*, *and patient arrival rate* λ_e^* *increase with* p.

From the proposition, we see that as the provider incentive p increases, structural quality investments and process quality both go up. The resulting increase in overall quality, combined with higher p, boosts the revenue earned per patient. Note that the increase in R_p could lead to an increase in the waiting time, which in turn could reduce the patient arrival rate ($\lambda_e = F(P) + \gamma R_s - aR_p$). However, the increase in the provider's capacity stemming from the structural quality investments more than compensates for the reduction in capacity resulting from the process quality improvement. Hence, the overall waiting time goes down, leading to an increase in the arrival rate as p goes up. In essence, the structural quality investments enable the provider to earn more per patient (due to increases in both structural and process quality) and also see more patients (due to the lower wait times that can attributed to an increase in the provider's capacity). Next, we look at the impact of changing the incentive offered to patients while keeping the provider incentive fixed.

Proposition 2. *For any given provider incentive* p *and combination of system parameters,* R_s^*, R_p^*, *and patient arrival rate* λ_e^* *increase with* P.

From Proposition 2, we see that as the patient incentive goes up, the provider increases both structural and process quality. The increase in process quality could result in longer waiting time for patients but this effect is countered by the increase in capacity from the structural quality investments and also a higher willingness to wait due to the increase in the patient incentive, and the net effect is that the patient arrival rate goes up. In essence, the increase in patient incentive results in more patients seeking care, and the provider responds by offering better quality of care since it leads to more revenues earned per patient over a larger patient pool. Having analyzed the provider's response to changes in incentives, we now turn our attention to the humanitarian organization's incentive design problem.

3.3. The Humanitarian Organization's Incentive Design Problem

The organization's objective (see problem IDP below) is to identify the right level of incentives to offer to patients and the provider in order to maximize quality-adjusted program coverage ($R\lambda_e$).

$$(IDP) \max_{p \geq 0, P \geq 0} \quad \Psi = R^*(p,P)\lambda_e^*(p,P),$$
$$\text{subject to} \quad \lambda_e^*(p,P)\left(pR^*(p,P) + P\right) \leq B.$$
$$\text{(Budget Constraint)}$$

Here B is the funding available to the humanitarian organization. We point out that in the above formulation, the objective function and the budget constraint are expressed on a per unit time basis. Also, notice that the above formulation does not exclude the possibility of offering both fixed and performance-based payments to the provider. If the contract has a fixed payment component, then B should be interpreted as the remaining funding available to offer incentives to patients and performance-based incentives to the provider.

As specified above, the humanitarian organization's objective is to maximize quality-adjusted program coverage, that is, $R^*(p,P) \times \lambda_e^*(p,P)$. From Proposition 1, we see that increasing the provider incentive p leads to an increase in both quality of care (R^*) and the patient arrival rate (λ_e^*). Proposition 2 indicates that increasing the patient incentive P also has a similar effect. Given these results, it is clear that the humanitarian organization would benefit from increasing both p and P. However, the limited funding availability may prevent the organization from increasing p and P simultaneously since offering more incentives to the provider results in less funding being available to offer incentives to patients (and vice-versa). Hence, to identify the right combination of p and P that maximizes the program impact, a budget-constrained organization needs to carefully evaluate the trade-off involved in increasing the incentive offered to the provider vs. the patients. We analyze this trade-off and present several key results regarding optimal incentives to patients and the health care provider in section 5. However, before looking at two-sided incentive design, we analyze one-sided incentive schemes that are frequently used in practice. In addition to being practically relevant, such one-sided incentive schemes also serve as a benchmark for the two-sided incentive design problem (IDP) presented earlier.

Prior to concluding this section, we comment on the payment scheme considered in our paper. Specifically, we address two key elements related to the payment scheme: 1. the linear nature of payments and 2. using aggregate quality scores for the payment instead of providing separate incentives based on the structural and process quality scores. We discuss and provide justification for both elements below.

Justification for the use of a linear payment scheme: Linear performance-based payment schemes have been used extensively in humanitarian health care programs, across a variety of health conditions (e.g., maternal and child health, primary care and HIV services) and countries (e.g., Rwanda, Nigeria, Senegal and Mozambique, see Gertler and Vermeersch 2012, USAID 2016a,b, and USAID 2016c). The main advantage of linear payment schemes is that they are relatively straightforward to understand and implement. Given the ease of implementation and relevance to practice, we focus attention on linear payment schemes in our work. A valid question that could arise in this context is whether there is a significant loss of performance due to the use of a linear payment scheme, relative to the optimal incentive scheme? To answer this question, we compared the quality-adjusted program coverage under the linear payment scheme considered in our paper with that under a centralized model where the humanitarian organization has full control over the health care provider and can directly make process and structural quality decisions. Our results indicate that the linear payment scheme compares favorably with the centralized model, with the performance gap being less than or equal to 2% in a overwhelming majority (nearly 81%) of the instances. We describe the centralized model in greater detail and present additional information regarding the performance gap in section 7.2.

Justification for using aggregate quality score in the payment scheme: In our work, we assume that the structural and process quality scores are combined into a single quality score R, when determining the payments to the provider. Against this background, one could argue that it might be better to specify separate payment rates to the provider based on the structural and process quality scores, instead of combining them into a single score. This is a valid concern but we justify the use of the aggregate quality score based on the following reasons: (1) In practice, the structural and process quality scores are typically combined into an aggregate measure R and there is extensive evidence (see section 3.2) that documents the widespread use of such incentives across a variety of humanitarian health care programs. (2) Making payments based on the individual quality components R_s and R_p instead of the aggregate score R does not lead to any significant improvements in program performance. For example, offering separate payments based on R_s and R_p results in an average performance improvement of roughly 1% over the payment scheme considered in this study that utilizes a combined quality score to determine provider payments. We present additional details regarding the payment scheme based on R_s and R_p in section 7.2 but the lack of any meaningful improvement is not surprising given that the payment scheme considered in our paper performs extremely well when stacked up against the centralized model (as mentioned in the previous paragraph). (3) Finally, having separate payment rates based on the individual quality components makes the model analytically intractable since the humanitarian organization now needs to make decisions regarding three incentives instead of two. This combined with points (1) and (2) raised above imply that a payment scheme based on the aggregate quality score is reasonable and well justified.

4. Analysis of the Current Practice: One-Sided Incentive Schemes

As highlighted earlier in the introduction, developing country health programs often focus exclusively on either the supply-side (incentives offered only to the provider, that is, $P = 0$) or the demand-side (incentives offered only to the patients, that is, $p = 0$)[3]. For example, pure supply-side incentive schemes are used in maternal and child health programs in Rwanda (Gertler and Vermeersch 2012), and also for a variety of health services (e.g., laboratory services, maternal and child health, and infectious diseases) in Burundi, Nigeria, Senegal and Mozambique (Rudasingwa et al. 2015, USAID 2016a,b, 2016c). Similarly pure demand-side incentive schemes are used in tuberculosis programs in Brazil, vaccination programs in India, and for maternal health in Bangladesh and Pakistan (see e.g., Carvalho et al. 2014, Jehan et al. 2012, Lagarde et al. 2007).

Given their widespread use in practice, we explore how the pure supply-side and pure demand-side incentive schemes perform relative to one another. Developing country health programs offer different types of services, and the relative importance of structural and process quality measures vary depending on the service offered. For instance, the health care programs in Rwanda and Mozambique assign higher weights to structural quality (i.e., $w_s > w_p$) for HIV services and lab testing (Gertler and Vermeersch 2012, USAID 2016c). In contrast, prenatal care services and family planning programs in Rwanda assign higher weights to process quality (i.e., $w_p > w_s$). Maternal care programs frequently assign roughly equal weights to process and structural quality measures. Since w_s and w_p values vary widely with the health care service offered, we are interested in understanding how the relative performance of the one-sided incentive schemes change across programs with varying emphasis on structural and process quality measures. In addition, we also explore how such one-sided incentive schemes stack-up against a two-sided incentive scheme where the humanitarian organization offers incentives to both patients and the health care provider (see section 5.2). We begin by analyzing one-sided incentive schemes.

4.1. Pure Demand-Side Incentive Schemes

Under a pure demand-side incentive scheme, the provider's earnings are not tied to either the quality of care or the number of people served (since $p = 0$). As a result, the provider is not motivated to invest in structural quality improvements, that is, $R_s^* = 0$. Furthermore, the provider is also indifferent to process quality choices since they do not require any direct material investments. In this case, we assume that the provider would be willing to set the process quality at $R_p^*(P) = \frac{F(P)}{2a}$, which is the socially desirable level from the perspective of the humanitarian organization. Under a pure demand-side incentive scheme, the organization's problem is the following.

$$\max_{P \geq 0} \quad \Psi = w_p R_p^*(P) \lambda_e^*(P),$$
$$\text{subject to } P\lambda_e^*(P) \leq B,$$

where $\lambda_e^*(P) = F(P) - aR_p^*(P)$. The above objective function is increasing in P. Hence, it is optimal to set P as high as possible subject to the budget constraint, that is, $P^* = \tilde{P}$ where $\tilde{P}$ solves $\tilde{P}\lambda_e^*(\tilde{P}) = B$. This observation immediately leads to the following result regarding the impact of budget.

COROLLARY 1. (ORGANIZATIONAL CHARACTERISTICS). *Under a pure demand-side incentive scheme, the patient incentive P^*, patient arrival rate λ_e^* and the quality of care R^*increase with B.*

The characterization of P^* preceding Corollary 1 is also useful in establishing how the program characteristics, namely the emphasis on structural and process quality measures, impact the provider's decisions and the incentive offered by the humanitarian organization.

PROPOSITION 3. (PROGRAM CHARACTERISTICS). *Under a pure demand-side incentive scheme, the provider's process and structural quality choices, and the incentive offered by the humanitarian organizations are invariant with the type of health care service offered (i.e., with w_s and w_p).*

The main implication of Proposition 3 is that under a pure demand-side incentive scheme, the provider is not motivated to take into consideration the attributes that are unique to each type of health care service when making the process and structural quality choices. Consequently, the patient incentive is also not tailored to the type of service offered. The fact that R_s^* is always zero under such a scheme could be especially problematic for health care programs that place a heavy emphasis on structural quality. As a result, we expect pure demand-side incentive schemes to perform poorly in programs where structural quality is highly important. We elaborate more on this in section 4.3. The last result in this section establishes the effect of the patient and provider characteristics.

PROPOSITION 4. *Under a pure demand-side incentive scheme, we have the following results regarding the impact of the different system parameters.*

1. ***Provider characteristics*** $(\boldsymbol{\alpha})$: *P^* decreases with α while λ_e^* and R^* increase.*
2. ***Patient characteristics*** $(\boldsymbol{V_b, C_t, c_w})$: *$P^*$ decreases with V_b while λ_e^* and R^* increase. The effect of C_t and c_w are opposite to that of V_b.*

To summarize, under pure demand-side incentives, we can fully characterize how the patient incentive changes with the program budget, nature of service offered, and patient and provider characteristics. Most of these results are on expected lines but interestingly, the patient incentive is not adapted to the specific type of health care service under consideration, that is, the humanitarian organization utilizes a one-size-fits-all approach across different types of programs. This directly impacts how the pure demand-side incentive scheme performs relative to supply-side focused incentives and also the two-sided incentive scheme. We discuss their relative performances in sections 4.3 and 5.2 respectively.

4.2. Pure Supply-Side Incentive Schemes

In this case, the humanitarian organization offers incentives only to the health care provider (i.e., $p = 0$). Under pure supply-side incentives, the provider's optimal quality choices (R_p^* and R_s^*) can be obtained from Lemma 1 by setting $p = 0$. These results form the basis for solving the humanitarian organization's incentive design problem specified below.

$$\max_{p \geq 0} \quad \Psi = R^*(p)\lambda_e^*(p),$$
$$\text{subject to} \quad pR^*(p)\lambda_e^*(p) \leq B,$$

where $\lambda_e^*(p) = F(0) + \gamma R_s^*(p) - aR_p^*(p)$. The above objective function increases (at least weakly) with p, which leads to the following characterization of the optimal provider incentive.

PROPOSITION 5. *Under a pure supply-side incentive scheme,*

1. *There exists a threshold $\tilde{B}$ such that for $B \leq \tilde{B}$, p^* is a solution of the budget constraint, that is, $p^*R^*(p)\lambda_e^*(p) = B$.*
2. *For $B > \tilde{B}$, $p^* = \frac{4a\beta}{w_p(a\frac{w_s}{w_p}+\gamma)(a\frac{w_s}{w_p}+\gamma+F(0))}$.*

The characterization of the threshold $\tilde{B}$ is provided in the proof of Proposition 5. The result below, derived from Proposition 5, establishes how the provider's optimal response and the incentive offered by the humanitarian organization change with the available budget.

COROLLARY 2. (ORGANIZATIONAL CHARACTERISTICS). *For budgets $B \leq \tilde{B}$, the provider payment p^*, quality of care offered R^* and the patient arrival rate λ_e^* all increase with B under pure supply-side incentives. For $B > \tilde{B}$, p^*, R^* and λ_e^* are all invariant with B.*

The fact that the p^* increases with B when the budget is below a threshold is not surprising. This is because the higher provider payment leads to better quality of care and also increases the patient arrival rate, both of which are beneficial to the humanitarian organization. However, having access to more funding is useful only when $B \leq \tilde{B}$. Beyond that, further increases in p^* do not translate into performance improvements since the provider's structural quality score is already at its peak value. When structural quality is at its peak, program performance can be increased further only through improvements in process quality and/or the patient arrival rate, both of which require offering incentives to patients. Unfortunately, the pure supply-side incentive scheme does not offer incentives to patients and as a result, the program performance flatlines for budgets beyond $\tilde{B}$, illustrating one of the main limitations of taking an exclusively supply-side focused approach. We shed more light on the drawbacks of using a pure supply-side focused incentive scheme as opposed to two-sided incentives in section 5.2. The following result highlights the impact of the program characteristics (i.e., weights w_s and w_p).

PROPOSITION 6. (PROGRAM CHARACTERISTICS). *Under a pure supply-side incentive scheme, ceteris paribus: (i) R_s^* increases with both w_s and w_p, (ii) R_p^* may increase or decrease with w_s while it always increases with w_p, and (iii) the humanitarian organization offers lower incentive (p^*) to the provider as either w_s or w_p increases.*

Proposition 6 demonstrates that consistent with expectations, the provider increases structural quality investments as w_s goes up. Under the pure supply-side incentive scheme, it can be shown that the overall quality score ($R^* = w_sR_s^* + w_pR_p^*$) and the arrival rate λ_e^* also increase with w_s, ceteris paribus. Thus, to balance the budget constraint, the humanitarian organization offers lower per-patient incentive to the provider as w_s increases.

It is interesting to note in Proposition 6 that the provider's structural quality investments increase not only with w_s but also with w_p. Note that structural quality investments lead to an increase in service capacity (by a factor $\gamma > 0$), which in turn allows the provider to spend more time with each patient, thereby increasing the process quality. This explains why structural quality investments increase with w_p. However, if $\gamma = 0$, structural quality investments do not increase the provider's capacity and in that case, we can show that the provider's structural quality investments *decrease* with w_p as one would expect. More broadly, our analysis indicates that whether or not the provider invests more in structural quality in response to changes in w_p depends on the extent to which structural quality investments impact the provider's capacity. When γ is less than a threshold, the relatively low benefits from the structural quality improvements (in terms of having additional capacity to offer better process quality) are outweighed by the costs. Hence, in such cases, the provider finds it beneficial to reduce structural quality investments as w_p goes up. However, for γ values above the threshold, the benefits of structural quality improvements are high and in those cases, the provider finds it beneficial to make additional investments in structural quality as w_p goes up since the additional capacity derived from those investments could be utilized to offer better process quality to patients. Given that the provider's response to w_p varies based on whether γ is low or high, the humanitarian organization also tailors the incentive offered accordingly – in settings where γ is high, the organization offers a lower per-patient incentive p as w_p goes up but this may not necessarily hold true in settings where γ is low. This suggests that even for similar type of health care services, what works in one setting might not necessarily work in others due to differences in the operating environment across settings.

A final observation we make from Proposition 6 is that under pure supply-side incentives, the provider's quality choices vary with w_s and w_p, which is different from the pure demand-side incentive scheme where the quality choices are indifferent to the nature of health care service offered by the organization. This has important implications for how the one-sided incentive schemes perform relative to one other, and how the relative performance changes based on the type of service offered. We discuss these issues in the next section. In the following proposition, we establish how the patient and provider characteristics impact the incentive offered and the overall program performance.

PROPOSITION 7. *Under a pure supply-side incentive scheme, we have the following results regarding the impact of the different system parameters.*

1. ***Provider characteristics*** $(\boldsymbol{\alpha})$*:* p^* *decreases with* α *while* λ_e^* *and* R^* *increase.*
2. ***Patient characteristics*** $(\boldsymbol{V_b}, \boldsymbol{C_t}, \boldsymbol{c_w})$*:* p^* *decreases with* V_b *while* λ_e^* *and* R^* *increase. The effect of* C_t *and* c_w *are opposite to that of* V_b*.*

In summary, the analysis of pure supply-side incentives highlights two key takeaways. First, the provider incentive and program performance increase with budget until a threshold but flatline beyond that. Hence, more funding may not necessarily translate into better performance under pure supply-side incentives. Second, humanitarian organizations need to be mindful of the nuances involved in tailoring the provider incentive to the type of service offered. For example, we see that the provider incentive p should be lowered as structural quality becomes more important. With respect to the emphasis on process quality, the humanitarian organization may find it optimal to decrease the provider incentive when γ is high but this result may not always hold true for low values of γ. This illustrates the need for humanitarian organizations to take into account the dependencies between program type and other factors in the operating environment when looking to tailor the provider incentive to the specific type of service offered by the program.

4.3. Performance Comparison of the One-Sided Incentive Schemes

In section 4.1, we noted that the provider does not make any structural quality investments (i.e., $R_s^* = 0$) under pure demand-side incentive schemes, regardless of the type of service offered by the program. However, under pure supply-side incentive schemes, the provider invests in structural quality improvements (i.e., $R_s^* > 0$, see Lemma 1 with $P = 0$) and furthermore, Proposition 6 establishes that R_s^* increases with w_s. For health care programs where structural quality is important, it is critical to create the right incentives for the provider to invest in structural quality improvements. In this context, the above discussion suggests that *pure supply-side incentive schemes are likely to perform better than pure demand-side incentive schemes in programs that place a high emphasis on structural quality.*

However, a similar result is not directly obvious for programs where process quality is deemed important. From section 4.1, we see that the process quality level under pure demand-side incentive schemes is $R_p^*(P) = \frac{F(P)}{2a}$, which, in theory, can be higher or lower than the process quality level under pure supply-side incentive schemes (follows from Lemma 1 with $P = 0$). However, the following proposition demonstrates that when the incentives are chosen optimally under each one-sided incentive scheme, *pure demand-side incentives perform better than pure supply-side incentives in programs that place a high emphasis on process quality.*

PROPOSITION 8. *For a given* w_p*, there exists a threshold value of* w_s*, say* $\tilde{w}_s$*, below which pure demand-side* *incentive schemes perform better than pure supply-side incentive schemes. The opposite is true for $w_s > \tilde{w}_s$.*

Proposition 8 highlights the importance of aligning the incentive scheme with the type of health care service offered by the humanitarian organization. In practice, both pure supply-side and pure demand-side incentive schemes are widely used but in many cases, the particular choice of the incentive scheme does not align well with the type of health care service offered. For example, pure supply-side incentive schemes have been used in prenatal care programs (a highly process-focused health care service based on weights w_s and w_p, see Gertler and Vermeersch 2012) in several countries but Proposition 8 suggests that humanitarian organizations could benefit by switching to pure demand-side incentive schemes for such programs. Of course, both the one-sided incentive schemes may not perform nearly as well as the two-sided incentive scheme but the above result suggests that even within the current practice of using one-sided incentives, there is significant scope for improvement (by as much as 14 times, on average, see Figure 2 in section 5.2) by ensuring that there is better alignment between the incentive scheme and the type of health care service offered.

5. Two-Sided Incentive Design: Analysis and Comparison with One-Sided Incentives

Having analyzed the one-sided incentive schemes, we now return to the two-sided incentive design problem (problem IDP) that we discussed earlier in section 3.3. We first present results related to the optimal incentives to patients and the health care provider. Then, we compare one-sided and two-sided incentive schemes.

5.1. Two-Sided Incentives: Insights

The following result characterizes the optimal patient and provider incentives in budget-constrained humanitarian health care programs.

PROPOSITION 9. *The optimal incentives to the provider and patients can be characterized as follows:*

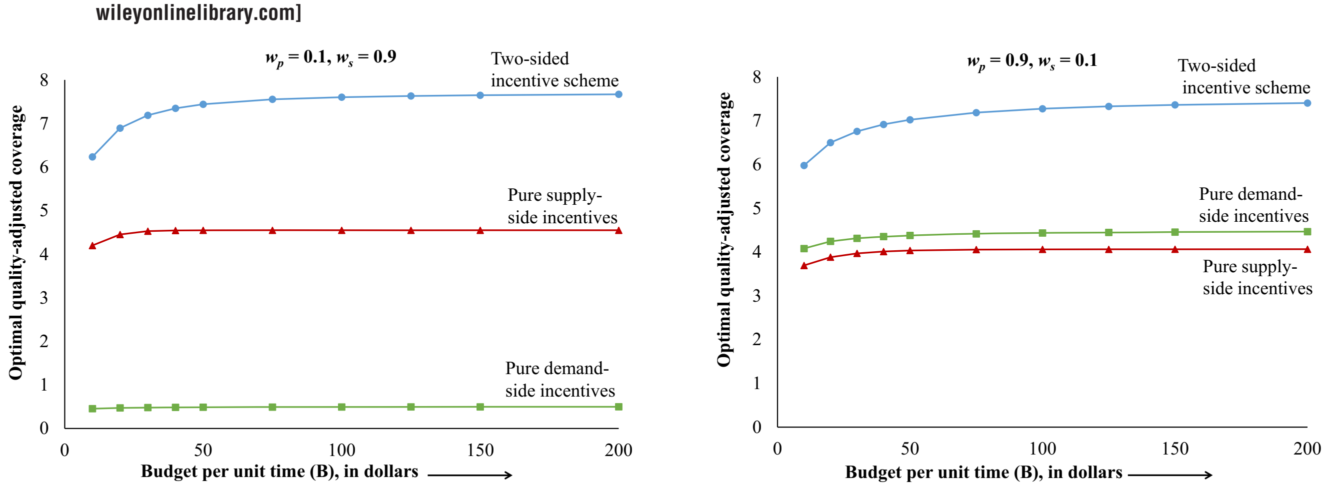
1. *There exists a threshold $\hat{B}$ such that for $B \le \hat{B}$, the budget B is allocated such that the marginal benefits of the provider incentive and patient incentive are the same. Specifically, there exists a unique solution $p^* > 0$ and $P^* > 0$ such that the budget constraint is binding and $\frac{\partial \Psi}{\partial p}\big|_{(p=p^*, P=P^*)} = \frac{\partial \Psi}{\partial P}\big|_{(p=p^*, P=P^*)}$, where Ψ is the humanitarian organization's objective function.*
2. *For $B > \hat{B}$, $p^* = \frac{4a\beta}{w_p(a\frac{w_s}{w_p}+\gamma)(F(P^*)+a\frac{w_s}{w_p}+\gamma))}$ and $P^* = \hat{P}(B)$, where $\hat{P}(B)$ is the solution of $\frac{1}{2}(F(P) + a\frac{w_s}{w_p} + \gamma)\left(\frac{2\beta}{a\frac{w_s}{w_p}+\gamma} + P\right) = B$.*

The following corollary, derived from Proposition 9, establishes how the incentives offered and program performance change with the humanitarian organization's budget.

COROLLARY 3. (ORGANIZATIONAL CHARACTERISTICS).

1. *For $B \le \hat{B}$, both incentives p^* and P^* increase with B. Furthermore, the quality of care offered R^* and the patient arrival rate λ_e^* also increase with B.*
2. *For $B > \hat{B}$, P^* increases with B while p^* decreases. However, both R^* and λ_e^* increase with B.*

Figure 2 Impact of Budget on the Quality-Adjusted Coverage Under Two-Sided and One-Sided Incentive Schemes [Color figure can be viewed at wileyonlinelibrary.com]



(a) (b)

Proposition 9 and Corollary 3 bring to light the nuances and challenges involved in designing incentives in budget-constrained humanitarian health care programs. The organization's objective is (at least weakly) increasing in both p and P and Proposition 9 establishes that for budgets less than the threshold $\hat{B}$, it is optimal to allocate the available funding such that the marginal benefits of the provider incentive and patient incentive are the same. If that is not the case, then the humanitarian organization can improve the quality-adjusted coverage by shifting funding away from the less-effective incentive type to the more-effective one until their marginal benefits become equal. This strategy of increasing both p and P is optimal until the provider's structural quality measure achieves its peak value, which happens at budget $\hat{B}$. However, once the budget exceeds $\hat{B}$, increasing p further leads to no improvement in either the arrival rate or the quality of care offered (since structural quality is already at its peak value and process quality is invariant to p when structural quality is at its peak). Hence, in this budget range, it is optimal for the humanitarian organization to combine an increase in the patient incentive with a reduction in the provider incentive. Such a strategy leads to an increase in the process quality and patient arrival rates while still maintaining the structural quality at its peak value.

The results presented in Proposition 9 are also useful in establishing how the incentives offered and the quality-adjusted coverage change based on the program characteristics. We summarize the results in the following proposition.

Proposition 10. (Program characteristics). *Under the two-sided incentive scheme, ceteris paribus:* (*i*) R_s^* *increases with both* w_s *and* w_p, (*ii*) R_p^* *may increase or decrease with* w_s *while it always increases with* w_p, *and* (*iii*) *With respect to the incentives, suppose budget B is such that* $B > \hat{B}$ *for a given* w_s *and* w_p. *Further suppose that* $B > \hat{B}$ *continues to hold for small enough increases in* w_s (*resp.,* w_p) *while keeping* w_p (*resp.,* w_s) *fixed. Then, the humanitarian organization offers a lower provider incentive* p^* *and a higher patient incentive* P^* *as either* w_s *or* w_p *increases.*

Similar to pure supply-side incentives, whether or not structural quality investments increase with w_p under two-sided incentives depends on the value of γ. Specifically, R_s^* increases with w_p when γ is above a threshold while it decreases with w_p otherwise. This reinforces our earlier comment regarding the importance of taking into consideration the specifics of the operating environment (for example, the impact of structural quality investments on the provider's capacity) when designing incentives to the health care provider as well as patients.

Propositions 9 and 10, concerning the effect of budget B and weights (w_s, w_p) on p^* and P^*, highlight the similarities as well as important differences between two-sided and one-sided incentive schemes. We elaborate more on this in the next subsection. We conclude this section by summarizing the impact of patient and provider characteristics under two-sided incentives.

Proposition 11. *Suppose budget B is such that* $B > \hat{B}$ *for a given set of problem parameters. Further suppose that* $B > \hat{B}$ *continues to hold for small enough increases in the parameter of interest while holding everything else constant. Then, we have the following results regarding the impact of the different system parameters under a two-sided incentive scheme.*

1. ***Provider characteristics*** $(\boldsymbol{\alpha})$***:*** *p^* and P^* decrease with α while λ_e^* and R^* increase.*
2. ***Patient characteristics*** $(\boldsymbol{V_b, C_t, c_w})$***:*** *$p^*$ and P^* decrease with V_b while λ_e^* and R^* increase. The effect of C_t and c_w are opposite to that of V_b.*

5.2. Comparison of the Two-Sided and One-Sided Incentive Schemes

In this section, we first compare the incentives offered by the humanitarian organization under the different incentive schemes before analyzing their relative performances.

5.2.1. Comparison of Incentives Offered. Comparing Corollaries 1 and 3, we see that the incentive offered to patients P^* always increases with B, implying that the effect of budget on patient incentives is directionally consistent across the one-sided and two-sided incentive schemes. However, the same cannot be said about the provider incentive p^*. Specifically, we see from Corollary 2 that under pure supply-side incentives, p^* increases with budget until a certain threshold but becomes flat beyond that. However, under the two-sided incentive scheme, p^* decreases with B beyond a threshold budget value while P^* simultaneously goes up (see Corollary 3). This strategy of pairing an increase in P^* with a reduction in p^* allows the humanitarian organization to make the most of the additional funding in this budget range. Unfortunately, the one-dimensional nature of the pure supply-side incentive scheme prevents the humanitarian organization from employing such creative strategies to improve program performance and consequently, the additional funding does not translate into performance improvement for budgets above a certain threshold. Another key implication of the above-mentioned results concerning p^* is that *the effects that one might intuitively expect to hold when considering the incentives in isolation (i.e., more funding* *translates into larger incentives) may not necessarily hold when multiple incentives are offered in combination*.

With respect to the program characteristics, we see that the effect of w_s and w_p on the provider incentive p^* is directionally similar across the one-sided and two-sided incentive schemes. However, that is not true in case of the patient incentive P^*. From Proposition 3, we see that under the pure-demand incentive scheme, P^* is invariant with the program type while the two-sided incentive scheme responds to changes in the program type by modifying the patient incentive accordingly (see Proposition 10). This enables the humanitarian organization to more precisely tailor the incentives offered to the specific program at hand instead of using a one-size-fits-all approach across different types of programs.

Finally, with respect to the patient and provider characteristics, our results indicate that their effect on the incentives is directionally similar under one-sided and two-sided incentive schemes. Overall, we see that there are similarities as well as key differences with respect to how the one-sided incentive schemes respond to changes in the system parameters, relative to the two-sided incentive scheme. It is important to note that even when the responses are directionally similar (e.g., p^* decreases with w_s under pure supply-side and two-sided incentive schemes), the actual incentives offered under the one-sided and two-sided incentive schemes can be quite different. This has important performance implications, which we discuss next.

5.2.2. Performance Comparison of the Two-Sided and One-Sided Incentive Schemes. In this section, we define performance gap as $\psi^* - \bar{\psi}$, where ψ^* is the quality-adjusted coverage under the two-sided incentive scheme and $\bar{\psi}$ is the quality-adjusted coverage under a one-sided incentive scheme. The following result establishes how the pure supply-side incentive scheme performs relative to the two-sided incentive scheme.

PROPOSITION 12. *The performance gap between the two-sided incentive scheme and the pure supply-side incentive scheme increases with B for budgets greater than a threshold.*

The next result characterizes the relative performance of the pure demand-side incentive scheme against the two-sided incentive scheme.

PROPOSITION 13. *Suppose $\gamma = 0$. Then,*

- *For exclusively process-focused programs (i.e., $w_s = 0$, $w_p > 0$), there is no performance gap between the two-sided and pure demand-side incentive schemes.*
- *For exclusively structure-focused programs (i.e., $w_s > 0$, $w_p = 0$), the performance gap between the two-sided and pure demand-side incentive schemes increases with B.*

Before getting into the conditions stated in Propositions 12 and 13, we highlight two major insights that emanate from these results. First, *there exist program types and operating environments (see Proposition 13) where offering incentives only to the patients is indeed optimal and does not lead to any performance loss when compared to the two-sided incentive scheme*. In contrast, offering incentives only to the health care provider is never optimal, regardless of the program type. The second key insight that we obtain from Propositions 12 and 13 is that the performance gap between the two-sided and one-sided incentive schemes increases with budget (with the exception of exclusively process-focused programs where the performance gap is zero as identified in Proposition 13). In the propositions, we demonstrate that this insight holds true under certain parametric conditions. For example, Proposition 12 requires the humanitarian organization's budget to be above a certain threshold. However, if this condition is not met, it is analytically intractable to show how the relative performance of the pure supply-side incentive scheme (relative to the two-sided incentive scheme) would change with budget. Similarly, in Proposition 13, we focus on two specific types of programs and also assume that $\gamma = 0$, that is, the structural quality investments do not lead to an increase in the provider's capacity. For more general program types and operating environments, it is difficult to analytically characterize how the relative performance of the pure demand-side incentive scheme changes with budget.

To that end, we conducted an extensive numerical study to see if the above-mentioned insights regarding the performance gap between the two-sided and one-sided incentive schemes also extend to more general settings. The results from the numerical study confirm that it is indeed the case, that is, the performance gap increases with budget. Specifically, we see from Figure 2a (resp. Figure 2b) that when using pure supply-side incentive schemes, the loss[4] in quality-adjusted coverage increases from roughly 28.4% (resp. 37.7%) at $B = 10$ to about 39.4% (resp. 45.5%) at $B = 200$ in case of health care programs that place a higher emphasis on structural quality (resp. health care programs that place a higher emphasis on process quality). We make a similar observation regarding the negative impact of using pure demand-side incentive schemes. *Taken together, these results highlight the importance of using the right incentive scheme for every organization but even more so for well-funded organizations with higher operating budgets*.

In addition to illustrating the performance gap, Figure 2 also provides several interesting insights into how the two-sided and one-sided incentive schemes work. To begin with, notice that the marginal benefit of additional funding approaches zero much earlier under one-sided incentives (beyond $B = 50$) when compared to the two-sided incentive scheme (beyond $B = 100$). This is true regardless of whether the program places a higher emphasis on structural or process quality. Furthermore, in regions where the quality-adjusted coverage increases with budget under both two-sided and one-sided incentive schemes ($B \leq 40$), the rate of increase is much smaller under one-sided incentives when compared to the two-sided incentive scheme. *Combined, the two results clearly highlight the significance of employing the right incentive scheme to fully realize the benefits of fundraising efforts by humanitarian organizations.*

A final observation here is that by switching from pure one-sided incentives to the two-sided incentive scheme, a higher quality-adjusted coverage can be achieved despite a lower budget. *This result assumes increased significance in today's aid environment where developing country health programs are exploring ways to reconcile the growing demand for health services against flatlining or shrinking aid funding received from developed nations* (Leach-Kemon et al. 2012).

In the next section, we present additional insights from our computational study, including how the performance gap between the one-sided and two-sided incentive schemes change with other factors in the operating environment including the provider's capacity and the nature of health care service offered by the humanitarian organization. Besides the provider's capacity and type of service offered, our computational study also offers insights regarding the effect of other system parameters, namely V_b (patient's perceived value of getting treatment), C_t (fixed cost of transportation to the health center), and c_w (patient's opportunity cost per unit time). However, most of those results are on expected lines and hence, we do not present them in the interest of space. We begin the next section by describing our test bed.

6. Computational Study

6.1. The Test Bed

We vary a number of parameters in our computational study and we use publicly available data sources to inform our parameter choices whenever possible. However, for some parameters, we do not have publicly available sources to guide the choice of parametric values and for those cases, we resort to exploring a wide range of possible values. In what follows, we discuss details regarding the values used in our computational study for two key parameters of interest, namely the program type and the humanitarian organization's budget.

Process and structural quality weights (w_p, w_s): Examples of programs where structural measures might be considered more important include HIV services and lab testing. For instance, the pay–for–performance program in Rwanda assigns weights of $w_s = 1$ and $w_p = 0$ for HIV and lab services (Gertler and Vermeersch 2012) while the program in Mozambique assigns weights of $w_s = 0.72$ and $w_p = 0.28$ for HIV services (USAID 2016c). In contrast, process quality measures might be considered more important in programs offering family planning and prenatal care services. For example, the pay-for-performance program in Rwanda assigns weights of $w_s = 0.12$ (resp. 0.22) and $w_p = 0.88$ (resp. 0.78) for prenatal care services (resp. family planning). Maternal care programs frequently assign roughly equal weights to process and structural quality measures. Inspired by these examples, we consider w_s and w_p values that add up to 1 but we emphasize that this assumption is not necessary for any of our results to hold. Consistent with the w_s and w_p values used in the above-mentioned health care programs, we consider two pairs of weights $(w_p, w_s) = (0.1, 0.9)$ and (0.3, 0.7) to represent a structure-focused program, $(w_p, w_s) = (0.9, 0.1)$ and (0.7, 0.3) to represent a process-focused program, and $(w_p, w_s) = (0.5, 0.5)$ to represent a program that places an equal emphasis on process and structural quality. In the interest of space, we present the results corresponding to $(w_p, w_s) = (0.1, 0.9)$ and (0.9, 0.1). The results are similar for other values of w_s and w_p.

Program budget *B*: The annual operating budget of health programs managed by humanitarian organizations and aid agencies exhibit a wide variation depending on a number of factors including the organization size, the catchment population size, and the nature of health care service offered. It is convenient to think of patient arrival rates and service rates on a per-hour or per-minute basis rather than annually and hence, we express the budget B also on a per-hour basis. To allow for a wide range of organization and program types, we consider the following ten values of B: 10, 20, 30, 40, 50, 75, 100, 125, 150, and 200. This translates to annual program budgets ranging from \$87600/year to about \$1.75 million/year. The budget range that we consider covers the operating budgets of several programs in practice ranging from the sexually–transmitted–infection (STI) program in Nicaragua (operating budget = \$60000/year) to reproductive health programs in Uganda (operating budget = \$2 million/year, see Witter and Somanathan 2012).

In addition to the above-mentioned parametric values, we also vary the following parameters in our study: provider's base capacity α (5 different levels), sensitivity of the provider's capacity to structural quality investments γ (3 levels), patients' willingness to wait captured by the ratio $\frac{V_b - C_t}{c_w}$ (5 levels), opportunity cost of waiting c_w (6 levels) and structural investment scaling factor β (3 levels). Details regarding the specific values considered for these parameters, along with a justification for their choices and data sources (if any), are provided in Appendix A. In total, considering the different combinations, we have 67,500 instances. In what follows, we discuss the insights from our computational study.

6.2. Impact of Program Type and Provider's Base Capacity Under the Two-Sided Incentive Scheme

Figure 3 demonstrates how the quality-adjusted coverage under the two-sided incentive scheme changes with the provider's base capacity (α) for different values of budget and program types. From the figure, we see that for any given value of B and program type, the quality-adjusted coverage increases with the provider's base capacity. More interestingly, notice that *the marginal benefit of additional funding increases with the provider's capacity*, with the additional benefits being roughly the same under structure-focused and process-focused programs. This observation highlights the critical role played by the provider's capacity in reaping the potential benefits of fundraising efforts. Having additional funding to offer incentives is more valuable when the system has adequate capacity to handle the additional influx of people.

Figures 4 and 5 illustrate how the patient and provider incentives change with the provider's base capacity for different program types. From the figures, we see that in both structure-focused and process-focused programs, the incentives offered decrease with the base capacity, reinforcing the analytical result presented in Proposition 11. However, the reduction is non-uniform, with larger reduction in incentives observed as we increase capacity from 2 to 6 than when capacity is increased from 6 to 10. This result is a direct consequence of how the waiting time changes in a non-linear fashion with the system capacity – as more capacity becomes available, the optimal incentives offered also change in a non-linear fashion to account for the reduction in the waiting times. A comparison of Figures 4a and b (or equivalently Figures 5a and b) also reveals that the sensitivity of incentives to the provider's capacity is higher in process-focused programs when compared to structure-focused programs. In practice, performance-based incentives change based on the program type but they are often not tailored to the characteristics of individual health care providers (see Gertler and Vermeersch 2012). The above results suggest that to achieve maximal impact, it is critical to factor in the provider's characteristics when designing incentives not only to the provider but to patients as well.

6.3. Impact of Program Type and Provider's Base Capacity On the Performance Gap Between Two-Sided and One-Sided Incentive Schemes

Figure 6 offers several insights into the performance of the two-sided and one-sided incentive schemes. From the figure, we see that *supply-side focused incentives perform better than pure demand-side incentives in structure-focused programs while the opposite is true (for the most part) for programs where process quality is considered more important.* The only exception is $\alpha = 2$ for process-focused programs where pure supply-side incentives perform better than demand-side focused incentives. In this case, the provider faces severe capacity constraints and using a strategy that focuses solely on the demand-side may not be effective in such situations since the provider does not have sufficient capacity to deal with the additional influx of people brought in through the incentives offered to patients. Overall, these observations are qualitatively consistent with Proposition 8, which suggests that supply-side focused incentives are likely to perform better than pure demand-side incentives in programs that place a high emphasis on structural quality (and vice-versa for programs that place a high emphasis on process quality). In addition to the relative performance of the one-sided incentive schemes, we also observe from Figure 6 that for both process-focused and structure-focused programs, there is a finite performance gap between the two-sided and one-sided incentive schemes.

Figure 6 provides additional insights regarding how the % loss in performance due to the use of one-sided incentive schemes vary depending on the program type and the provider's base capacity. From the figure, we observe that the negative impact of using pure supply-side incentive schemes (relative to the two-sided incentive scheme) increases with the provider's capacity while the negative impact of using pure demand-side incentive schemes decreases with capacity. This has different implications depending on whether a program is structure-focused or process-focused. In case of structure-focused programs, *the difference in % loss under the two one-sided incentive schemes narrows down with capacity while the opposite is true for process-focused programs*. Intuitively, as the base capacity increases, the humanitarian organization may find it beneficial to dedicate more resources to lower the demand-side barriers in order to encourage people to seek care. As a result, the effectiveness of pure demand-side incentive schemes increases with capacity while the opposite is true for supply-side focused incentives. This, combined with our earlier result that supply-side focused incentives perform

Figure 3 Impact of the Provider's Base Capacity on Quality-Adjusted Coverage in (a) Structure-Focused ($w_p = 0.1, w_s = 0.9$) and (b) Process-Focused ($w_p = 0.9, w_s = 0.1$) Programs [Color figure can be viewed at wileyonlinelibrary.com]

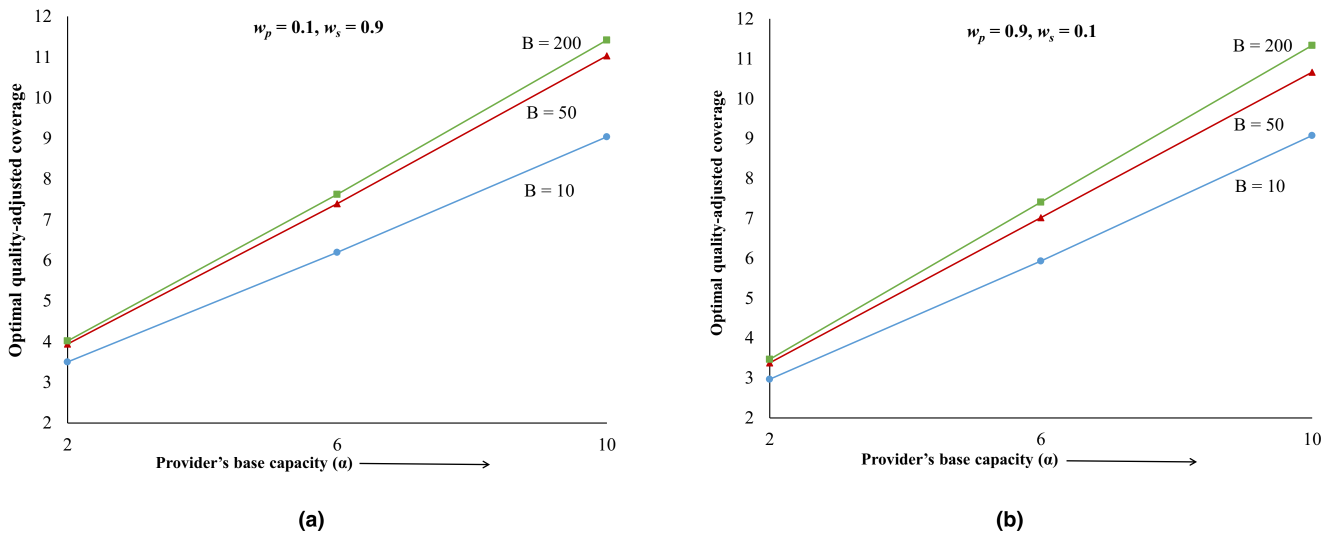
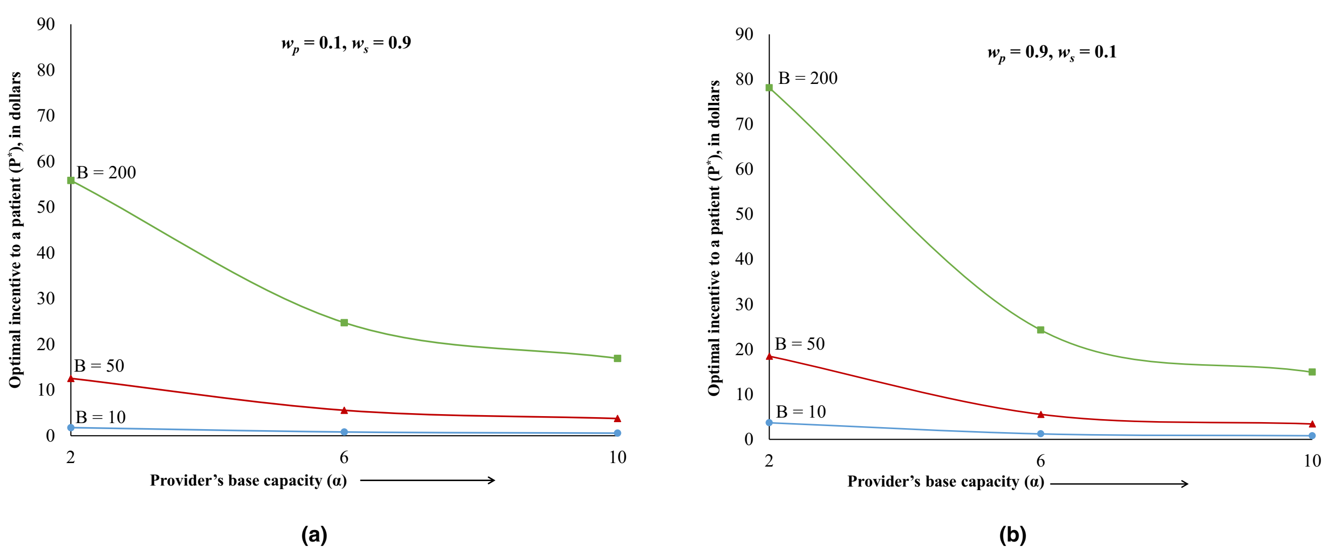


Figure 4 Impact of Provider's Base Capacity on the Incentive Offered to Patients in (a) Structure-Focused ($w_p = 0.1, w_s = 0.9$) and (b) Process-Focused ($w_p = 0.9, w_s = 0.1$) Programs [Color figure can be viewed at wileyonlinelibrary.com]



better than pure demand-side incentives in structure-focused programs, explains why the difference in % loss under the two one-sided incentive schemes reduces with capacity in Figure 6a. Notice, however, that even when this difference is at its lowest value, the % loss under pure demand-side incentives is almost double the % loss under supply-side focused incentives. Hence, humanitarian organizations that currently use pure demand-side incentives in structure-focused programs can significantly improve performance by switching to supply-side focused incentives. Making this switch is certainly a valuable first step but we wish to point out that using a more effective one-sided incentive scheme does not get humanitarian organizations close to achieving the full potential in case of structure-focused programs (the minimum % loss under pure supply-side incentives is roughly 17% and increases further as α goes up). To truly make the best use of available resources, the organization needs to complement the performance-based incentives to the provider with incentives to patients in order to simultaneously increase the quality of care offered and the number of patients seeking access to care.

A similar insight also emerges for process-focused programs, that is, humanitarian organizations can substantially improve performance by switching to a more effective one-sided incentive scheme but they are still likely to under-perform significantly, relative to the two-sided incentive scheme (% loss in

Figure 5 Impact of Provider's Base Capacity on the Incentive Offered to the Provider in (a) Structure-Focused ($w_p = 0.1, w_s = 0.9$) and (b) Process-Focused ($w_p = 0.9, w_s = 0.1$) Programs [Color figure can be viewed at wileyonlinelibrary.com]

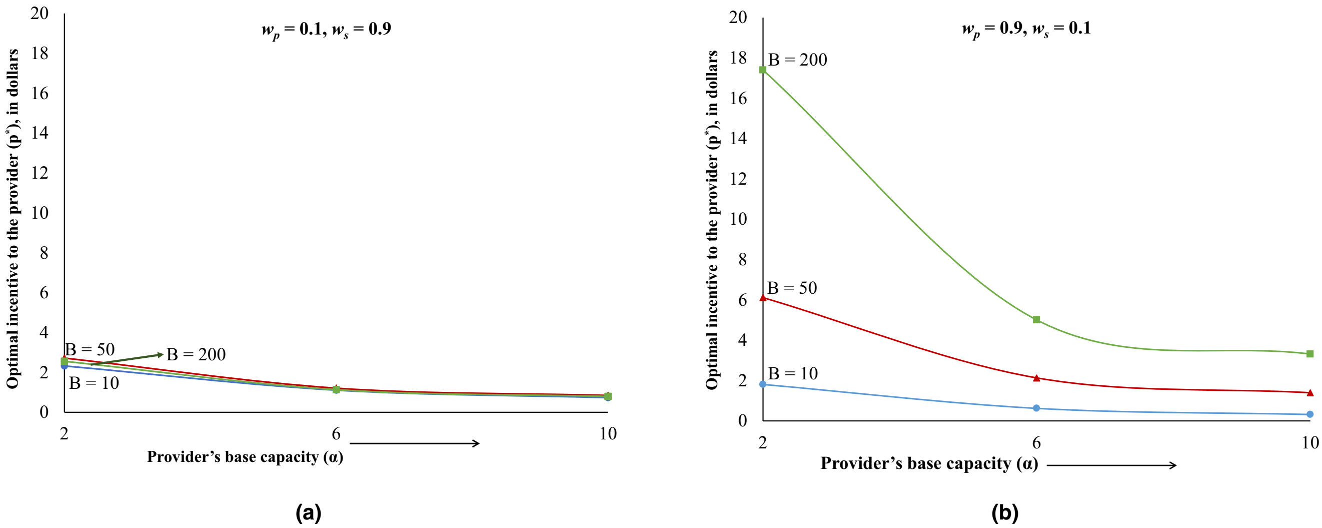
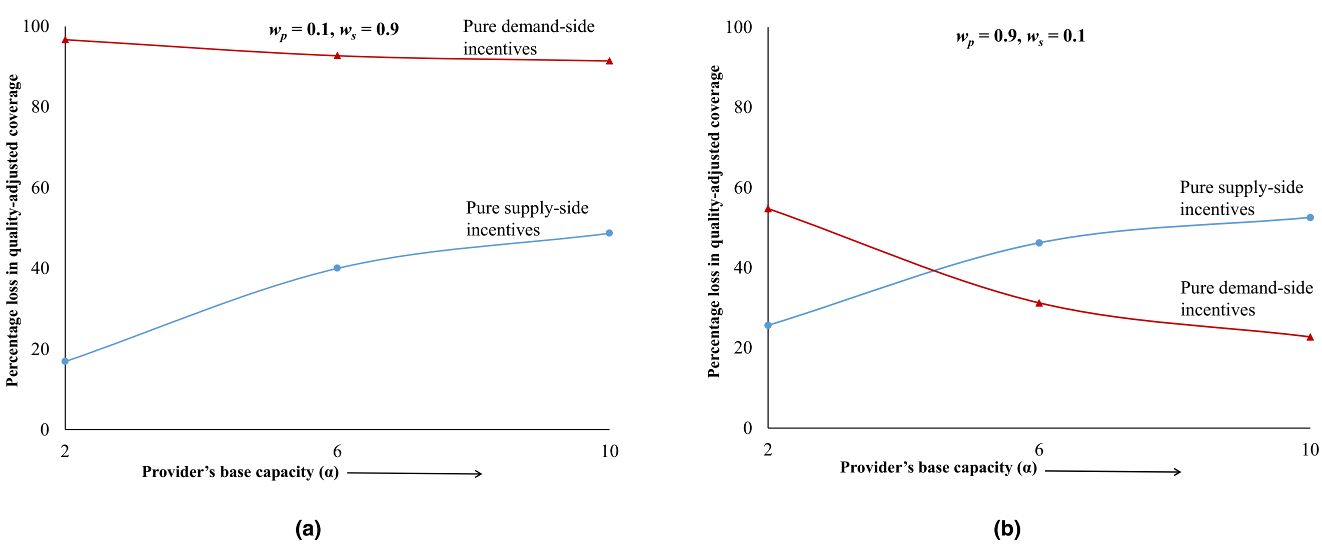


Figure 6 Percentage Loss in Quality-Adjusted Coverage Under One-Sided Incentives (Relative to the Two-Sided Incentive Scheme) in (a) Structure-Focused ($w_p = 0.1, w_s = 0.9$) and (b) Process-Focused ($w_p = 0.9, w_s = 0.1$) Programs [Color figure can be viewed at wileyonlinelibrary.com]



performance is 25% or more in process-focused programs). This highlights the urgent need for humanitarian organizations to move away from their current practice of using one-sided incentive schemes to a more balanced approach that carefully combines patient and provider incentives.

7. Model Extension and Justification for the Aggregate-Quality Based Payment Scheme

In this section, we first consider a more general version of our model to test the robustness of our findings and to ensure that they are not driven by the specific modeling assumptions. Subsequently, we consider alternate payment schemes to the health care provider (including a centralized model where the humanitarian organization and the provider are part of the same entity) and present evidence to demonstrate that the use of a linear, aggregate-quality based payment scheme is well-justified and does not lead to any noticeable loss of performance relative to potentially more complex payment schemes.

7.1. Model Extension

In the model described in section 3, we assume that a patient's decision to seek or to not seek care depends on the waiting time, which is indirectly influenced by the provider's process and structural quality choices. However, we do not *explicitly* account for the impact of process and structural quality on the patients' care-seeking behavior. USAID (2016c) mentions that patients consider the availability of staff, provider competence, and wait times when making their care-seeking decisions, suggesting that the quality of care may not directly enter into the patients' decisions in developing country settings (as long as the provider is competent enough, which we assume they are). However, according to Kiguli et al. (2009), factors affecting the perceived quality of services and thus, utilization, may include "long waiting times, poor geographical access, poor infrastructure and hygiene, lack of equipment and qualified staff, and interpersonal interactions." This implies that along with waiting time and geographical access (factors captured in our model), patients may also care about process and structural quality when deciding whether or not to seek care. In our model, the humanitarian organization works to ensure that patients receive a certain level of quality of care through quality-adjusted payments but to keep the analysis tractable, we have thus far assumed that the provider's quality choices do not *directly* influence the patient arrival rate. We relax this assumption in the model extension presented below.

Recall that V_b represents a patient's perceived value of getting treatment, which was assumed to be independent of process and structural quality. We relax this assumption by modeling the perceived valuation to be linearly increasing in both structural and process quality scores. Let $\mathcal{V}_b$ denote the modified perceived value of receiving treatment. Then, we have $\mathcal{V}_b = V_b + \theta_s R_s + \theta_p R_p$, where $\theta_s, \theta_p \geq 0$. Thus, the effective arrival rate of patients to the health center is given by $\lambda_e = \mu - \frac{c_w}{\mathcal{V}_b + P - C_t}$.

The above change in the perceived value of treatment, while seemingly minor, makes the model intractable and precludes obtaining analytical results. Hence, we resort to a computational analysis of the model using a representative subset (see Table B1 in Appendix B) of the test bed discussed in section 6. The results from this analysis indicate that all the key insights (related to quality-adjusted coverage and the patient and provider incentives) obtained from our current model continue to hold under the extended setting as well, demonstrating their robustness.

7.2. Justification for the Aggregate-Quality Based Payment Scheme

We begin this section by considering a centralized model where the humanitarian organization and the health care provider are part of the same entity. The analysis of the centralized model serves as a theoretical benchmark to evaluate the "goodness" of the solutions obtained assuming a linear, aggregate-quality based payment scheme. In addition, it is also useful in understanding how much of the performance gap between the centralized model and the aggregate-quality based payment scheme can be bridged by utilizing potentially more complex payment schemes.

7.2.1. Centralized Model. Under a centralized model, the humanitarian organization and health care provider are assumed to be part of the same entity and as such, the organization has full control over the provider's structural and process quality decisions. In addition, consistent with our original model, the humanitarian organization offers incentive P to patients to encourage them to seek care. The humanitarian organization's problem under a centralized setting is the following.

$$\max_{P\geq 0,\ 0\leq R_s,R_p\leq 1} \Psi = \left(w_s R_s + w_p R_p\right)\left(F(P) + \gamma R_s - aR_p\right),$$
$$\text{subject to } P\left(F(P) + \gamma R_s - aR_p\right) + \beta R_s^2 \leq B.$$

Notice that the total payment to the patients and the structural quality investments, combined, cannot exceed the budget available. For the centralized model, we can analytically characterize the optimal patient incentive, the optimal process quality level and the amount of funding to utilize toward structural quality improvements. We do not present those results in the interest of space.

Table 2 provides a summary of the numerical analysis we conducted to evaluate the performance of the linear, aggregate-quality based payment scheme considered in our paper, relative to the centralized model. From the table, we see that the aggregated-quality based payment scheme stacks up relatively well against the centralized model, with the performance gap being less than or equal to 1% (resp. 2%) in nearly 76% (resp. 81%) of the instances. The aggregate-quality-based payment scheme considered in our paper is highly relevant to practice (mirroring the payment schemes used in several humanitarian health care programs in developing countries) and the results in Table 2 offer further justification for the use of such a scheme by demonstrating that there is little to no loss of performance relative to the centralized model. An additional implication of this result is that switching from the current payment scheme to an alternate one is not likely to lead to any noticeable improvements in performance. We elaborate more on this in the next section.

7.2.2. Offering Separate Payments Based on R_s and R_p. In the model described in section 3, the provider payment is based on the aggregate quality score. However, since the humanitarian organization has full information about the structural and process

Table 2 Performance Gap between the Centralized Model and the Linear, Aggregate-Quality Based Payment Scheme

Gap	≤0.5%	≤1%	≤2%	≤5%	≤10%
Percentage of instances	71%	76%	81%	86%	91%

quality measures, one could argue that it might be better to offer separate payments based on structural and process quality instead of combining them into one payment. For example, for every patient treated, the provider could be offered a payment of p_1 (resp. p_2) for unit improvements in structural quality (resp. process quality). In this case, the provider's problem is as follows:

$$\max_{0 \le R_p, R_s \le 1} \pi = \left(p_1 R_s + p_2 R_p\right)\left(F(P) + \gamma R_s - a R_p\right) - \beta R_s^2.$$

The humanitarian organization's problem under this setting is as follows:

$$\max_{p_1, p_2, P \ge 0} \quad \Psi = (w_s R_s^* + w_p R_p^*)\left(F(P) + \gamma R_s^* - a R_p^*\right),$$
$$\text{subject to} \quad \lambda_e^*\left(p_1 R_s^* + p_2 R_p^* + P\right) \le B.$$

Notice that in the above formulation, the organization makes decisions regarding three incentives, not two as in our original model. This makes the formulation analytically intractable. Hence, we resort to a numerical study to analyze the benefits of offering separate payments based on R_s and R_p. Our results indicate that offering payments based on the individual quality components does not lead to any significant improvement in program performance, with the average improvement being only 1% over the aggregate-quality based payment. This is not entirely surprising in light of the results discussed in the previous section – since the aggregate-quality based payment performs well relative to the centralized model (and by extension, to the unknown optimal payment scheme), it is only natural that the switch to a different payment scheme did not result in a considerable improvement in performance, further justifying the payment scheme considered in our paper.

8. Conclusion

In this study, we focus on humanitarian health care programs in developing countries and analyse how a budget-constrained organization should design incentives to patients and the health care provider in order to maximize a program's impact. Identifying the right combination of incentives assumes increased significance in light of the underwhelming effectiveness of pure supply-side focused approaches in several cases and also the fact that the impact of demand-side barriers are more pronounced in developing country settings. In response to demand-side barriers, humanitarian organizations have begun providing incentives to patients. Nevertheless, the incentive schemes used in practice continue to be one-sided, focusing primarily on either the supply-side or the demand side.

Our work sheds light onto the value of combining the patient and provider incentives as opposed to the one-sided incentive schemes that are frequently used in developing country health programs. Recall from section 4 that pure supply-side incentive schemes have been used in structure-focused programs in countries like Rwanda and Mozambique and they have been used in process-focused programs in Rwanda and Burundi. Our results indicate that by moving from the pure supply-side incentive scheme to two-sided incentives, humanitarian organizations can improve program performance by 78%, on average, in structure-focused programs and by 129%, on average, in process-focused programs. Pure demand-side incentives have been primarily used in process-focused programs and for such programs, using two-sided incentives can lead to a 103% improvement (on average) in program performance. Our analysis also indicates that the performance gap between one-sided and two-sided incentive schemes grows with budget. Hence, using the right incentive scheme is important for every organization but even more so for well-funded organizations with higher operating budgets.

In addition to illustrating the value of combining patient and provider incentives, our work also brings forth several insights regarding how budget-constrained organizations should structure the incentives to achieve maximal impact. We demonstrate that it is always beneficial to increase the incentive offered to patients if additional funds become available. However, it may not be always optimal to increase the incentives offered to the provider and patients *simultaneously* – we find that in certain settings, the optimal strategy is to pair the increase in patient incentive with a reduction in the incentive offered to the provider.

Finally, our analysis identifies several key factors that could impact the potential benefits derived from fundraising efforts by humanitarian organizations. One such factor is the provider's capacity. Our results suggest that having additional funds to offer incentives is valuable only when the provider has sufficient capacity to handle the additional influx of people. We also find that fundraising efforts will have little to no impact on program performance unless the correct incentives are in place for patients to seek care and for providers to offer the best possible care.

We conclude this study by briefly discussing some of our modeling assumptions as it relates to contract theory. For example, we assume that the provider faces limited liability, that is, the payments to the provider have to be non-negative. One question that could come up in this context is whether relaxing this assumption would impact our results. Payment schemes involving penalties are rarely, if ever, used in humanitarian health care programs and our analysis indicates that relaxing the limited liability assumption would not materially alter our results since the payment scheme considered in this study performs well relative to the centralized model presented in section 7.2.1 (performance gap is less than 2%). Another aspect that one could explore is the possibility of information asymmetry between the humanitarian organization and the provider as it relates to the provider's "type." Here type could refer to the sensitivity of the provider's capacity to the structural quality investments (which, for simplicity, could be either high or low). Given the information asymmetry, the humanitarian organization can design a menu contract that offers different per-patient payment rates targeted at the different provider types. Under such a scenario, we expect the effectiveness of the supply-side incentives offered to the provider to go down relative to the setting where there is no information asymmetry. The implication of this result is that for programs where pure supply-side incentives perform better than demand-side focused incentives under no information asymmetry, the performance gap between the two may narrow down in the presence of information asymmetry. On the contrary, for programs where pure demand-side incentives perform better than supply-side focused incentives, the performance gap may widen when there is information asymmetry. In a similar vein, we expect the benefits of using a two-sided incentive scheme (over the pure demand-side incentive scheme) to be lower in the presence of information asymmetry. Future research that rigorously analyzes the impact of information asymmetry could shed additional light on these dimensions.

Acknowledgments

The authors are grateful to the Department Editor, the Senior Editor and two anonymous referees for their constructive feedback, which has helped in improving the contents and exposition of this study significantly. In addition, the authors thank the seminar participants at Washington University and the conference participants at the 2017 POMS Annual Conference, and the 2016 and 2018 INFORMS Annual Meeting for their valuable feedback.

Appendix A. Test Bed

In this section, we provide details regarding the parametric values used in our computational study. These are in addition to the parameters discussed in section 6.1.

Provider's base capacity α: The parameter α captures the provider's base capacity, that is, the average number of people served per hour when $R_s = 0$ and $R_p = 0$. In our study, we consider five values of α between 2 and 10 in increments of 2.

Sensitivity of the provider's capacity γ: The parameter γ determines the sensitivity of the provider's capacity to structural quality investments. We consider the following three values of γ: 0, 2.5 and 5. This ensures that we have γ values that are both less than and greater than $a\frac{w_s}{w_p}$, relaxing the assumption $\gamma \geq a\frac{w_s}{w_p}$ that we specified in section 3.

Patients' willingness to wait: Recall from section 3 that the effective arrival rate of patients (λ_e) can be identified by solving the following equation: $\frac{V_b+P-C_t}{c_w} = \frac{1}{\mu-\lambda_e}$. The ratio $\frac{V_b+P-C_t}{c_w}$ has a natural and intuitive interpretation: it captures a patient's willingness to wait for treatment (in time units), a key factor that impacts a patient's decision to seek or to not seek medical care. For a patient, the best case scenario is when he/she does not wait and proceeds directly to receive treatment – in this scenario, the expected time spent by the patient in the system will be $\frac{1}{\alpha+\gamma-aR_p}$. In our work, we assume that the ratio $\frac{V_b-C_t}{c_w} \geq \frac{1}{\alpha-a}$ – this ensures that at least in the scenario where patients don't have to wait, they would be willing to seek care even in the absence of any incentives (p = 0). Moreover, by choosing $\frac{V_b-C_t}{c_w}$ to be at least $\frac{1}{\alpha-a}$, we ensure that the entire range of process and structural quality choices are available to the provider. For our study, we consider five values of the ratio $\left(\frac{V_b-C_t}{c_w}\right)/\left(\frac{1}{\alpha-a}\right)$: 1.1, 1.25, 1.5, 1.75, and 2. We set a=1 throughout the study. Recall that the α values used in our study vary between 2 and 10 – this relaxes the assumption $\alpha < 2a$ that we mentioned in section 3.

Opportunity cost c_w: In our model, c_w represents the patient's opportunity cost per unit time associated with waiting and receiving care. In many cases, people take time off from work or activities like farming to seek medical care and hence, we use the minimum wage data published by the International Labour Organization (http://www.ilo.org/ilostat/GWR) as a guiding principle to determine the appropriate value of c_w. According to this data, the minimum wage in developing and under-developed countries varies between \$0.01/hour to \$5/hour. To capture this range, we consider the following six values of c_w in dollars per hour: 0.1, 0.25, 0.5, 1, 2.5, and 5.

Structural investment scaling factor β: As discussed in section 3, structural measures capture the physical and infrastructure-related aspects of providing care like facility cleanliness, availability of appropriate health technologies and qualified personnel, management capabilities and so on. Accordingly, the structural quality investments could range from employing more staff to procuring new medical equipment to providing training in financial and general management skills. To be consistent with other parameters, the structural quality investments are represented in \$/hr – capital investments to buy equipment and training-related costs are assumed to be amortized. In our study, for each c_w, we varied β so that the ratio β/c_w took on values 1, 3, and 5, respectively. This approach allows us to capture different types of structural improvement investments – at the low end ($\beta/c_w = 1$ with $c_w = 0.1$), the annual investment required to improve the structural quality score from 0 to 1 is \$876. This is comparable to the cost of hiring an additional employee in certain under-developed countries to perhaps improve facility cleanliness or assist the physician in their day-to-day activities at the clinic. On the higher end ($\beta/c_w = 5$ with $c_w = 5$), the annual investment to improve R_s from 0 to 1 is \$219000. This is consistent with large capital investments required to procure medical equipment.

Appendix B. Test Bed: Model Extension

Table B1 Test Bed for the Model Extension

Parameter	Values
w_p	0.1, 0.5, 0.9
B (in dollars)	10, 50, 200
γ	0,2.5, 5
α	2, 6, 10
$\left(\frac{V_b - C_t}{c_w}\right)/\left(\frac{1}{\alpha - a}\right)$	1.5
c_w	2.5
β/c_w	3
a	1
θ_s	1, 2
θ_p	0.5, 1.5

The above test bed is used to check if the key insights obtained from our main model described in section 3 also continue to hold under the model extension presented in section 7.1.

Notes

[1]We use the term "humanitarian" organization broadly in this study to refer to different types of organizations including public health agencies, local non-profit organizations as well as in-country operations of global health agencies (for example, UNICEF, USAID etc.).

[2]R_s^*, R_p^* and the patient arrival rate λ_e^* are functions of p and P. However, for ease of exposition, we do not explicitly state the dependence whenever there is no notational confusion.

[3]We wish to emphasize that $p = 0$ does not imply that the total payment to the provider is zero. In fact, payment schemes with $p = 0$ are frequently used in several developing country health programs where the provider is paid a fixed salary and the compensation is not tied to performance.

[4]% loss $= \frac{\psi^* - \bar{\psi}}{\psi^*}$

References

Alizamir, S., F. Iravani, H. Mamani. 2019. An analysis of price vs. revenue protection: Government subsidies in the agriculture industry. *Management Sci.* **65**(1): 32–49.

Amudhan, S., K. Mani, S. K. Rai, C. S. Pandav, A. Krishnan. 2013. Effectiveness of demand and supply side interventions in promoting institutional deliveries: A quasi-experimental trial from rural north India. *Int. J. Epidemiol.* **42**(3): 769–780.

Anand, K. S., M. P. Fazil, S. Veeraraghavan. 2011. Quality speed conundrum: Trade-offs in customer-intensive services. *Management Sci.* **57**(1): 40–56.

Arora, P., M. Rahmani, K. Ramachandran. 2017. Balancing advisory and service delivery efforts in nonprofits: A service design perspective. Working paper, Scheller College of Business, Atlanta, GA.

Berenguer, G., Q. Feng, J. G. Shanthikumar, L. Xu. 2017. The effects of subsidies on increasing consumption through for-profit and not-for-profit newsvendors. *Prod. Oper. Manag.* **26** (6): 1191–1206.

Boyce, N. 1996. Counterpoint: Using outcome data to measure quality in health care. *Int. J. Qual. Health Care* **8**(2): 101–104.

Carvalho, N., N. Thacker, S. S. Gupta, J. A. Salomon. 2014. More evidence on the impact of India's conditional cash transfer program, Janani Suraksha Yojana: Quasi–experimental evaluation of the effects on childhood immunization and other reproductive and child health outcomes. *PLoS ONE* **9**(10): e109311.

Chang A. Y., K. Cowling, A. E. Micah, A. Chapin, C. S. Chen, G. Ikilezi, N. Sadat, G. Tsakalos, J. Wu, T. Younker, Y. Zhao. 2019. Past, present, and future of global health financing: A review of development assistance, government, out-of-pocket, and other private spending on health for 195 countries, 1995-2050. *Lancet* **393**(10187): 2233–2260.

Chen, J., A. Qi, M. Dawande. 2019. Supplier centrality and auditing priority in socially–responsible supply chains. *Manuf. Serv. Oper. Manag.* Forthcoming.

Dai, T., S. Cho, F. Zhang. 2016. Contracting for on-time delivery in the U.S. Influenza vaccine supply chain. *Manuf. Serv. Oper. Manag.* **18**(3): 332–346.

Donabedian, A. 2005. Evaluating the quality of medical care. *Milbank Q* **83**(4): 691–729.

Ensor, T., S. Cooper. 2004. Overcoming barriers to health service access: Influencing the demand side. *Health Policy Planning* **19** (2): 69–79.

Fofanah, M. 2010. SIERRA LEONE: Unfulfilled Promise of Free Maternal Health Care for Mothers. Available at http://www.ipsnews.net/2010/10/sierra-leone-unfulfilled-promise-of-free-maternal-health-care-for-mothers/ (accessed date October 12, 2019).

Gallien, J., I. Rashkova, R. Atun, P. Yadav. 2017. National drug stockout risks and the Global Fund disbursement process for procurement. *Prod. Oper. Manag.* **26**(6): 997–1014.

Gandjour, A. 2010. Investment in quality improvement: How to maximize the return. *Health Econ.* **19**(1): 31–42.

Gertler, P., C. Vermeersch. 2012. Using performance incentives to improve health outcomes. Working paper, World Bank, Washington, DC.

Grainger, C., A. Gorter, J. Okal, B. Bellows. 2014. Lessons from sexual and reproductive health voucher program design and function: A comprehensive review. *Int. J. Equity Health* **13**(1): 33.

Hayford, T. B., J. L. Maeda. 2017. Issues and challenges in measuring and improving the quality of health care. Working paper, Congressional Budget Office, Washington, DC.

Heese, H. S., J. M. Swaminathan. 2006. Product line design with component commonality and cost-reduction effort. *Manuf. Serv. Oper. Manag.* **8**(2): 206–219.

Jehan, K., K. Sidney, H. Smith, A. de Costa. 2012. Improving access to maternity services: An overview of cash transfer and voucher schemes in South Asia. *Reprod. Health Matters* **20** (39): 142–154.

Jonasson, J. O., S. Deo, J. Gallien. 2017. Improving HIV early infant diagnosis supply chains in Sub-Saharan Africa: Models and application to Mozambique. *Oper. Res.* **65**(6): 1479–1493.

Kiguli, J., E. Ekirapa-Kiracho, O. Okui, A. Mutebi, H. MacGregor, G. W. Pariyo. 2009. Increasing access to quality health care for the poor: Community perceptions on quality care in Uganda. *Patient Prefer. Adherence* **3**: 77.

Kostami, K., S. Rajagopalan. 2014. Speed quality trade-offs in a dynamic model. *Manuf. Serv. Oper. Manag.* **16**(1): 104–118.

Lagarde, M., A. Haines, N. Palmer. 2007. Conditional cash transfers for improving uptake of health interventions in low- and middle-income countries: A systematic review. *JAMA* **298**(16): 1900–1910.

Leach-Kemon, K., D. P. Chou, M. T. Schneider, A. Tardif, J. L. Dieleman, B. PC. Brooks, M. Hanlon, C. JL. Murray. 2012. The global financial crisis has led to a slowdown in growth of funding to improve health in many developing countries. *Health Aff.* **31**(1): 228–235.

Mangham, L. J., K. Hanson. 2010. Scaling up in international health: What are the key issues? *Health Policy Plan.* **25**(2): 85–96.

Mu, L., M. Dawande, X. Geng, V. Mookerjee. 2016. Milking the quality test: Improving the milk supply chain under competing collection intermediaries. *Management Sci.* **62**(5): 1259–1277.

Natarajan, K. V., J. M. Swaminathan. 2014. Inventory management in humanitarian operations: Impact of amount, schedule, and uncertainty in funding. *Manuf. Serv. Oper. Manag.* **16**(4): 595–603.

Natarajan, K. V., J. M. Swaminathan. 2019. Coordinating demand and supply in funding–constrained developing country health supply chains. Working paper, Carlson School of Management, Minneapolis, MN.

Paul, A., T. Rajapakshe, S. Mallik. 2019. Socially optimal contracting between a regional blood bank and hospitals. *Prod. Oper. Manag.* **28**(4): 908–932.

Privett, N., F. Erhun. 2011. Efficient funding: Auditing in the nonprofit sector. *Manuf. Serv. Oper. Manag.* **13**(4): 471–488.

Rudasingwa, M., R. Soeters, M. Bossuyt. 2015. The effect of performance–based financial incentives on improving health care provision in Burundi: A controlled cohort study. *Glob. J. Health Sci.* **7**(3): 15–29.

Silverman, R., M. Over, S. Bauhoff. 2015. Aligning incentives, accelerating impact: Next generation financing models for global health. Report, Center for Global Development, Washington, DC.

Taylor, T. A., W. Xiao. 2014. Subsidizing the distribution channel: Donor funding to improve the availability of malaria drugs. *Manage. Sci.* **60**(10): 2461–2477.

The AIDSTAR-2 Project. 2011. The PBF handbook: designing and implementing effective performance–based financing programs. Available at https://www.msh.org/resources/the-pbf-handbook-designing-and-#implementing-effective-performance-basedfinancing-programs (accessed date October 12, 2019).

The World Bank. 2015. Compare MDG progress status across groups. Available at http://data.worldbank.org/mdgs/progress-status-across-groups-number-of-#countries (accessed date October 12, 2019).

USAID. 2016a. Quality of care in performance–based incentives programs: Nigeria case study. Report, USAID, Washington, DC.

USAID. 2016b. Quality of care in performance–based incentives programs: Senegal case study. Report, USAID, Washington, DC.

USAID. 2016c. Quality of care in performance–based incentives programs: Mozambique case study. Report, USAID, Washington, DC.

Wang, X., Q. Wu, G. Lai, A. Scheller-Wolf. 2019. Offering discretionary healthcare services with medical consumption. *Prod. Oper. Manag.* **28**(9): 2291–2304. https://doi.org/10.1111/poms.13051.

Witter, S., A. Somanathan. 2012. Demand–side financing for sexual and reproductive health services in low and middle–income countries: a review of the evidence. Working paper, World Bank, East Asia and the Pacific Region.

World Bank Group. 2015. Achieving results for women's and children's health. Available at https://www.rbfhealth.org/publication/achieving-results-women%E2%80%99s-and-childrens-health-2015-progress-report (accessed date October 12, 2019).

Yamey, G. 2012. What are the barriers to scaling up health interventions in low and middle income countries? A qualitative study of academic leaders in implementation science. *Glob. Health* **8**(1): 11.

Yang, Y., J. V. D. Broeck, L. M. Wein. 2013. Ready-to-use food-allocation policy to reduce the effects of childhood undernutrition in developing countries. *PNAS* **110**(12): 4545–4550.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1: Proofs