

Corinth, Kevin C.

Working Paper

Designing performance-based incentives when service providers compete for users to help

AEI Economics Working Paper, No. 2016-10

Provided in Cooperation with:

American Enterprise Institute (AEI), Washington, DC

Suggested Citation: Corinth, Kevin C. (2016) : Designing performance-based incentives when service providers compete for users to help, AEI Economics Working Paper, No. 2016-10, American Enterprise Institute (AEI), Washington, DC

This Version is available at:

<https://hdl.handle.net/10419/280546>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



AMERICAN ENTERPRISE INSTITUTE

Designing Performance-Based Incentives when Service Providers Compete for Users to Help

Kevin Corinth
American Enterprise Institute

AEI Economics Working Paper 2016-10
June 2016

© 2016 by the American Enterprise Institute for Public Policy Research. All rights reserved.

The American Enterprise Institute for Public Policy Research (AEI) is a nonpartisan, nonprofit, 501(c)(3) educational organization and does not take institutional positions on any issues. The views expressed here are those of the author(s).

Designing Performance-Based Incentives when Service Providers Compete for Users to Help

Kevin Corinth*
American Enterprise Institute

June 1, 2016

Abstract

Social service providers such as teachers, healthcare providers and homeless shelters receive trillions of dollars each year to help people. Recently, policymakers and other funders have attempted to obtain better outcomes by implementing performance-based incentive schemes that pay more money to higher performing providers. In this paper, I develop a simple model of this incentive design problem with a distinguishing feature—providers compete for users to help by adjusting service quality. I characterize a broad class of incentive schemes that elicit efficient service quality, and I show that popular incentive schemes (including value-added and pay-for-percentile schemes) are generally suboptimal and can have perverse distributional consequences, even when all characteristics of individuals are observed. I discuss implications for performance-based incentives in education and healthcare.

*This paper is based on a chapter from my dissertation entitled “Competition and Homelessness.” I thank Tony Cookson, Elizabeth Corinth, Ryan Dorow, Steve Levitt, Jens Ludwig, Casey Mulligan, Tobias Salz, and Caroline Weber for very helpful comments. Please email any comments to the author at kevin.corinth@aei.org.

1 Introduction

Governments and individuals in the United States spend trillions of dollars each year funding social service providers, including teachers, physicians and homeless shelters. Dissatisfaction continues to mount, however, regarding the performance of providers in spending resources efficiently: Spending on public education totals \$600 billion, while just 35% of eighth-grade students are proficient in math. Spending on healthcare totals \$2.8 trillion, but costs vary dramatically across the country with no associated difference in outcomes achieved.¹ Aside from concerns about aggregate performance, many observers worry that the most disadvantaged users receive especially low service quality. In an attempt to boost overall achievement and curtail distributional inequities, funders have in recent years begun implementing performance-based incentives for service providers. Schools are punished for failing to meet basic standards, teachers receive bonuses for better than expected performance under value-added payment schemes, and surgeons receive good marks on report cards for low mortality rates. These incentive schemes have been studied extensively and are often criticized for encouraging improvements at only certain parts of the achievement distribution, measuring achievement poorly, or failing to fully account for characteristics of users which affect performance.² If achievement could be measured perfectly and the incentive designer had full information regarding user characteristics, however, it is believed that some of the existing incentive schemes could perform efficiently.³

What each of these performance-based incentive schemes ignore, and what this paper highlights, is that service providers may compete for users by adjusting service quality. When users with certain characteristics are more profitable to service providers than others,

¹The education figures come from U.S. Census Bureau (2013) and National Center for Education Statistics (2013); the healthcare figure comes from Centers for Medicare and Medicaid Services (2013).

²For example, Neal and Schanzenbach (2010) find that threshold-based incentives in No Child Left Behind legislation increased test scores only at the middle of the achievement distribution; Stecher (2002) reviews evidence of catering instruction to tests independent of knowledge formation; and Dranove et al. (2003) discusses the impact of not accounting for all patient characteristics in surgeon report cards.

³See Neal (2011) and Barlevy and Neal (2012) for a discussion of some of these issues, as well as an incentive scheme—“pay-for-percentile”—which does not even require perfect measurement of achievement, but rather, a rank-preserving measure of it.

providers will try to attract more profitable users by increasing effort or spending—resources that could have been more effectively spent on others. This form of competition obviously has perverse consequences for performance-based incentives which do not account for user characteristics; for example, teachers will expend disproportionate effort on students with naturally higher achievement levels. Less obviously, competition for users also undermines popular incentive schemes which fully account for user characteristics. Consider, for instance, value-added incentive schemes, which are used in bonus payments for teachers and report cards for surgeons. Value-added schemes reward user-performance which exceeds the predicted performance level for users with the same characteristics. When providers compete for users, these schemes will elicit identical service quality across all users, rather than eliciting higher quality service to those it helps the most. Conversely, a scheme which simply pays providers a per-user fee which exactly compensates for each user’s *optimal* service quality level, completely ignoring user performance and actual service quality provided, is optimal. In this case, providers compete for users by bidding up service quality until the fee exactly compensates for their effort. These examples illustrate that while competition for users undermines popular performance-based incentive schemes, it can also serve to replace them.⁴

This paper considers a simple incentive design problem from the perspective of a third-party funder who wishes to help other people (users) achieve specific outcomes. I assume that there are many service providers and that heterogeneous users can choose to utilize the service provider that offers them the highest service quality. Service providers, meanwhile, maximize profit (or utility) by adjusting service quality in order to attract more profitable users, with competition driving profits to zero. I show that under popular incentive schemes, competition increases service quality to users for whom it is not most effective. As a result,

⁴This contrasts with the recommendation in Neal (2011), who suggests that teacher incentive schemes are best viewed as complements to policies which promote student choice. Note, however, that performance-based incentives can be used as complements to competition if designed appropriately. For instance, efficient pay-for-percentile schemes in Barlevy and Neal (2012), which vary base payments depending on user characteristics in order to exactly satisfy provider participation constraints, can still elicit efficient service quality across users in the presence of competition.

these schemes are generally suboptimal and have undesirable distributional consequences. I also characterize the set of optimal incentive schemes.

The paper contributes to the literature on performance based incentives by highlighting the importance of competition for users in designing efficient incentive schemes. Prior research on the design of performance incentives has focused on a number of other important challenges, including risk aversion of agents (Holmstrom 1979; Holmstrom and Milgrom 1987), multiple tasks (Holmstrom and Milgrom 1991), the inability to contract on outcomes of primary interest (Baker 1992), and the inability to contract on cardinal measurements of outcomes (Barlevy and Neal 2012). However, these studies do not allow the service quality of providers to affect the set of users with whom they are matched. Within the healthcare literature, meanwhile, it is established that competition for patients can affect the efficiency of different payment systems (Feldstein 1971; Joskow 1980; Held and Pauly 1983; Ma and McGuire 1997; Ellis 1998; Kessler and McClellan 2000). However, such models do not generally consider the design of performance incentives given the difficulty in measuring healthcare outcomes. By allowing service providers to compete for users on the basis of service quality, this paper incorporates a potentially important feature of social service markets not examined in the previous literature on the design of performance-based incentives. The simple model considered here should be extended in future research to more complex models that match particular real-world settings.

The rest of the paper proceeds as follows: Section 2 presents the model, characterizes optimal incentive schemes, and evaluates several other popular incentive schemes. Section 3 discusses applications of the model to specific social service sectors—school teachers and physicians. Section 4 concludes.

2 The Model

This section first describes the model and proves several results. Then it analyzes how several real-world incentive schemes perform in the context of the model, highlighting both aggregate efficiency and distributional consequences.

2.1 The Model and Results

There are three types of actors considered in the model—(i) users, (ii) service providers, and (iii) the funder. Given some incentive scheme designed by the funder, service providers maximize profit or utility by choosing some service quality level and some set of users to serve. Users are free to choose the service provider that offers them the highest service quality. In the model, I both characterize the set of optimal incentive schemes and show that popular incentive schemes are not contained in this set.

First consider the users to potentially be helped (e.g., students, patients, homeless individuals, etc.). Suppose there are I users of T types. Denote $w_i \in \mathbb{R}$ as the outcome achieved by user i . The outcome might represent human capital—expressed as a level or change—for students, survival for surgery patients, or achievement of housing for the homeless.⁵ Each type t has an expected production function $f_t(s)$ which maps service quality s (e.g., dollars or effort expended on the user) into expected production of the outcome, where $f_t(\cdot)$ is continuous and twice differentiable, $f'_t(\cdot) > 0$, and $f''_t(\cdot) < 0$. Let e_i denote a stochastic productivity shock equal to the difference between actual production and expected production, so that $w_i = f_{t_i}(s) + e_i$ for a user i who receives service quality s , and $E[e_i|t = t_i] = 0$. Assume that users prefer higher values of the outcome, or equivalently since f is increasing, greater s to less. A user who is not helped receives the outside option equal to $f_t(0)$.

Next consider the service providers (e.g., teachers, physicians, homeless shelters, etc.) who provide services to users to help them achieve outcomes. Assume there are J identical providers, and that $J > T$. Each provider j chooses some level of per-user service quality s_j ,

⁵In the case of survival and achievement of housing it will be convenient to use $w_i \in \{0, 1\}$.

and some set of users $\Psi_j \subset I$. Also, assume that each user helped by provider j receives the same service quality s_j .⁶ Provider objectives will be made explicit following development of notation for incentive schemes.

The final player in the helping market is a funder (e.g., government or a representative donor) which determines an incentive scheme—some way to fund providers—on the basis of the users they serve and the outcomes their users achieve. The funder cannot, however, observe service quality.⁷ Assume that the reward (or price) paid to providers is additively separable over users helped and nondecreasing in outcomes achieved. Additionally, service providers are assumed to be price-takers in that the types of users a given provider serves and the outcomes they achieve have no effect on the reward granted *for any given* set of users served and the outcomes they achieve. Price-taking is realistic if there are a large number of service providers. Formally, all of this means that the total reward P , granted to a service provider j , can be written as

$$P_j = \sum_{i \in \Psi_j} p(t_i, w_i; \{t_{i'}, w_{i'}\}_{i' \notin \Psi_j}) \quad (1)$$

I will generally write the user specific reward as $p(t_i, w_i)$, though it should be remembered that the reward may depend on user types and outcomes achieved at all other providers. Also, I will refer to the function p as the incentive scheme chosen by the funder.

Given any incentive scheme, provider objectives can now be specified. In particular, service providers are assumed to maximize a payoff function

$$\pi(s_j, \Psi_j) = \sum_{i \in \Psi_j} (E[p(t_i, w_i)] - s_j) \quad (2)$$

The payoff function considered here can represent different classes of provider objectives.

⁶The alternative, that providers are allowed to choose different service quality for different users, would not change the results except to require fewer providers to guarantee the zero profit condition.

⁷The assumption that funders cannot observe service quality does not change the results regarding competition and characterization of optimal incentive schemes. It simply rules out an incentive scheme that pays providers only if they provide the optimal level of service quality.

If p is monetary reward and s_j is spending, then the payoff function represents expected profit. If s_j is disutility of effort, meanwhile, the payoff function represents expected utility, where p may now also represent nonmonetary rewards.⁸ With the model notation and setup specified, I now define and characterize an equilibrium.

Definition 1. *Fix some incentive scheme p . The set of service quality levels and user assignments $\{s_j^*, \Psi_j^*\}_{j \in J}$ is an equilibrium if it solves*

$$\begin{aligned} \max_{s_j, \Psi_j} \pi &= \sum_{i \in \Psi_j} (E[p(t_i, w_i)] - s_j) \\ \text{s.t. } s_j &\geq s_k^* \text{ if } \Psi_j \cap \Psi_k^* \neq \emptyset \quad \forall k \in J \end{aligned} \quad (3)$$

for all $j \in J$, and $\Psi_j^* \cap \Psi_k^* = \emptyset$ for all $j, k \in J$.

In equilibrium, each provider maximizes payoff by choosing some service level s_j^* and some set of users Ψ_j^* such that it beats the offers of all other providers that help the same user; the condition $\Psi_j^* \cap \Psi_k^* = \emptyset$ is required to maintain unique assignments in the case of providers offering the same service level. Having defined equilibrium, there are two additional assumptions necessary to deal with “ties” and indifference. First, if users are offered the same, highest offer of service quality by multiple providers, I will assume that they divide up equally among these providers. Second, if a provider offers a higher service quality level than another provider, I will assume that the higher quality provider is willing to accept users from the lower quality provider as long as doing so does not decrease his payoff.

I can now state results regarding equilibrium payoff and strategies of providers. I first prove a lemma that no provider can receive a higher payoff than any other in equilibrium.

Lemma 1. *All service providers have the same payoff in equilibrium, that is, $\pi_j^* = \pi_k^* \forall j, k \in J$.*

Proof. See Appendix. □

⁸Note that these alternative interpretations of the payoff function require particular assumptions on the utility function.

Essentially, if any provider had a strictly higher payoff than any other, then another provider could increase its own payoff by mimicking the choice of users and incrementally increasing service quality over the first provider. Using this result that payoffs must be equal across all service providers in equilibrium, I can state the main results regarding strategies of the providers. Proposition 1 states that in equilibrium, each provider must have zero expected payoff, and Corollary 1 states that all users of a given type are treated equally.

Proposition 1. *In equilibrium, the expected payoff for each service provider is zero, that is, $\pi_j^* = 0 \forall j \in J$.*

Proof. See Appendix. □

Corollary 1. *All users of a given type receive the same equilibrium service quality, that is, $s_j^* = s_k^*$ if $t_i = t_{i'}$ and $i \in \Psi_j^*$, $i' \in \Psi_k^*$.*

Proof. See Appendix. □

That payoffs are zero across all providers follows from the assumption of a greater number of providers than user types—there must exist some user type which multiple service providers serve, leading these providers to compete payoffs down to zero. And as a result of Lemma 1, all providers must therefore receive zero payoff. The corollary that all users of a given type receive the same service quality follows directly from zero equilibrium payoff for all providers, since violation of equal treatment would necessarily allow for strictly positive payoff.

Having derived equilibrium allocations for any given incentive scheme, I can now characterize the funder’s optimal scheme.⁹ Define an optimal incentive scheme as follows:

Definition 2. *p is called an optimal incentive scheme for (expected) budget M and value*

⁹Note that for schemes representing nonmonetary preferences or for which multiple sources of funding exist, the scheme may not be a choice.

weights $\{\lambda_t\}_{t \in T}$ if it solves

$$\begin{aligned} \max_{\{P_j\}_{j \in J}} \quad & \sum_{i \in I} \lambda_{t_i} w_i \\ \text{s.t.} \quad & E \left[\sum_{i \in I} p(t_i, w_i) \right] \leq M \end{aligned} \tag{4}$$

The following proposition fully characterizes the set of optimal incentive schemes; it simply states that payoff must be zero at the optimal service level and negative for all higher levels of service quality.

Proposition 2. *p is an optimal incentive scheme if and only if*

1. $E[p(t, f_t(s_t^E))] = s_t^E$, and
2. $E[p(t, f_t(s))] < s$ for all $s > s_t^E$

for all $t \in T$, where $\{s_t^E\}_{t \in T}$ is given by the set of T equations $\lambda_1 f_1'(s_1^E) = \lambda_t f_t'(s_t^E)$ for $t > 1$ and $\sum_{i \in I} s_{t_i}^E = M$.

Proof. See Appendix. □

Incentive schemes of many forms are contained in this optimal set. Perhaps the simplest is one for which $p(t, w) = s_t^E$ for all $w \geq 0$, that is, a fixed fee equal to the optimal service quality level – paid to the provider regardless of outcomes achieved. This scheme completely ignores performance, and the only monitoring required is of the number of users served of each type. It elicits optimal service levels because providers must compete for users, bidding up service quality until the payoff is driven down to zero. For the set of optimal schemes in general, the price paid for outcomes below the optimal level is completely irrelevant (given the assumption of nondecreasing price), and the price paid for outcomes above the optimal level is irrelevant as long as payoff is negative. Figure 1 illustrates some examples of optimal schemes.

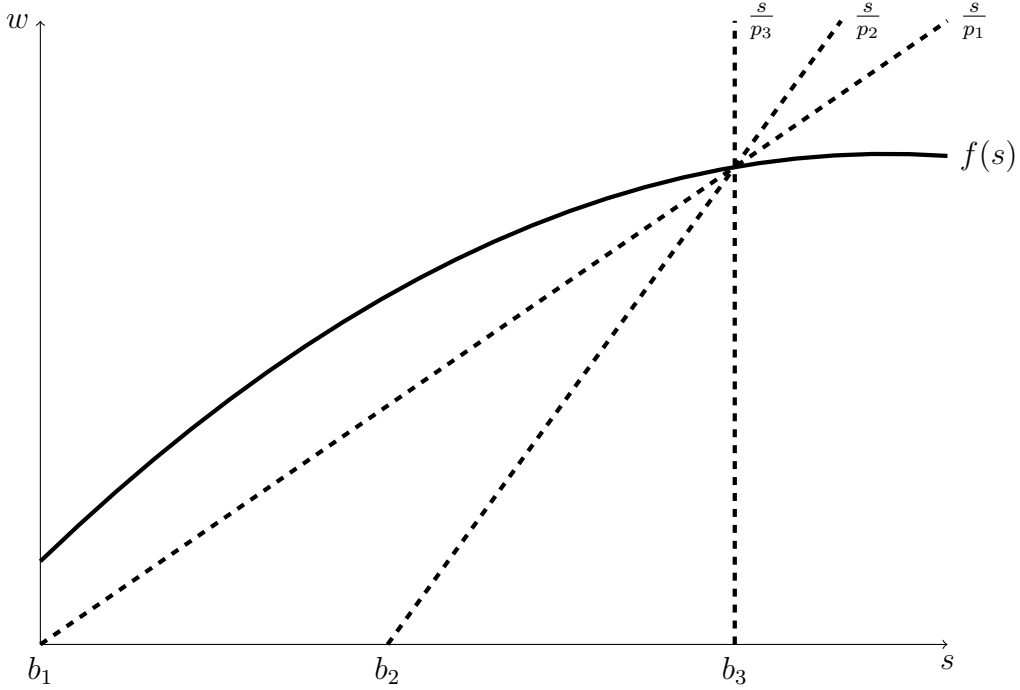


Figure 1: Optimal Incentive Schemes

The following three schemes all lead to the funder's optimal solution, where b_3 is the optimal service level. For pricing system i , b_i denotes the unconditional fee paid to a provider for serving the user, and p_i denotes the additional reward paid per unit of outcome achieved by the user. Note that an infinite number of linear as well as nonlinear schemes achieve the optimal solution as well.

2.2 Incentive Scheme Examples

I next consider the performance of incentive schemes commonly used in practice, in terms of both aggregate outcomes achieved and distributional effects on different user types. In order to compare them to optimal incentive schemes, it will in some cases be useful to compare users with the same marginal productivity of service quality, but with different levels of productivity.

Definition 3. Suppose $f_r(s) = f_t(s) + \alpha$ for some $\alpha > 0$. Then call type r easier-to-help than type t , and call type t harder-to-help than type r . Also, an incentive scheme is said to underserve hard-to-help users if (i) $f_t(s) = f(s) + \alpha_t$ for all $t \in T$, and (ii) $s_r^* \geq s_t^*$ if type r is easier to help than type t for all $r, t \in T$, with one inequality strict.

It should be clear that if a user can be called easier-to-help than another user, both users

will receive the same service quality in any optimal scheme where their value weights are equal. Also, in some proofs I will assume that types can be characterized in this way to show a case in which a scheme is suboptimal, but there will generally be many other production functions which lead to suboptimal allocations of service quality. The first scheme I consider is a level-based incentive scheme, which has been proposed (and frequently criticized) as a method of rewarding teachers.

Definition 4. *A level-based incentive scheme has the property that*

$$p(t, w; \{t_i, w_i\}_{i \notin \Psi_j}) = \begin{cases} x_l & \text{if } w < \bar{w} \\ x_h & \text{if } w \geq \bar{w} \end{cases}$$

where $x_l < x_h$ and $\bar{w} > 0$.

Proposition 3. *A level-based incentive scheme is suboptimal, and underserves hard-to-help users.*

Proof. See Appendix. □

That level-based incentive schemes underserve the lowest achieving users is hardly surprising; perhaps less obvious, however, is that users who would achieve the standard regardless of service quality are not underserved. Providers must still compete for high-achieving users, and so while provider service quality does not improve the likelihood of the user achieving the standard, it is required to induce the user to choose the provider. Figure 2 illustrates a level-based incentive scheme.

I next consider an outcome-based incentive scheme which pays providers on the basis of type-specific outcomes achieved, with a pure outcome-based scheme that ignores types as a special case. The pure outcome-based scheme might be a good representation of individual donations to charities which must compete for users to serve. These funding decisions may not explicitly compare performance relative to that of other providers, but rather, pay on the basis of absolute performance.

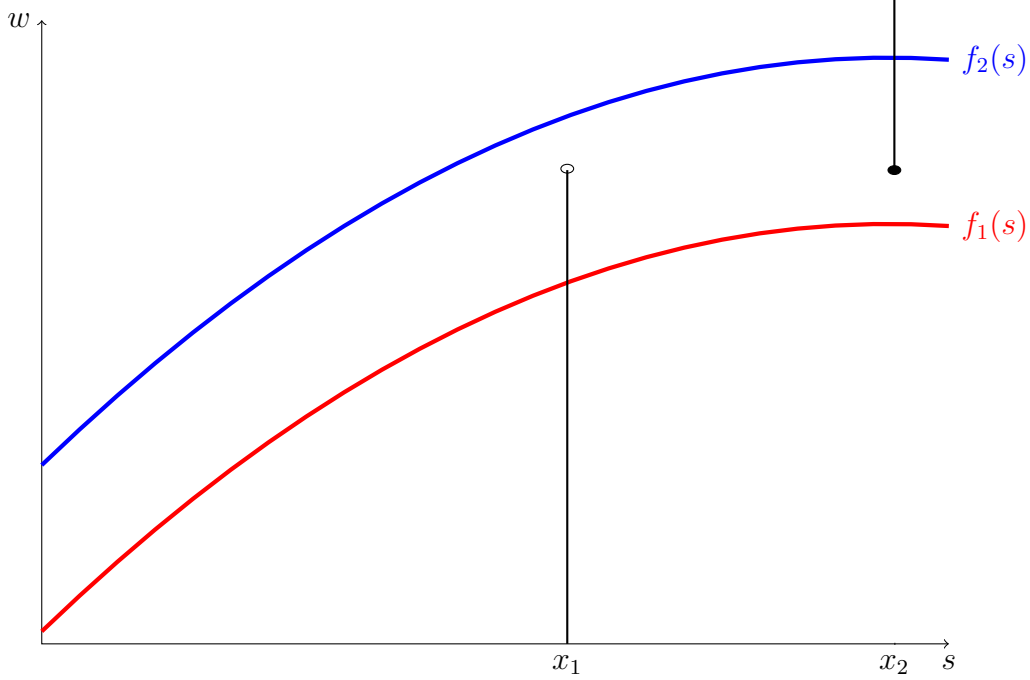


Figure 2: Level-Based Incentive Schemes

A level-based incentive scheme leads to overspending on the easier-to-help type relative to the harder-to-help type. Type 1 receives x_1 and type 2 receives x_2 in service quality.

Definition 5. *An outcome-based incentive scheme has the property that*

$$p(t, w; \{t_i, w_i\}_{i \notin \Psi_j}) = h(t, w)$$

A pure outcome-based incentive scheme has $h(t, w) = h(w)$ for all $t \in T$.

Proposition 4. *An outcome-based incentive scheme can be optimal if designed appropriately.*

A pure outcome-based incentive scheme is suboptimal and underserves hard-to-help users.

Proof. See Appendix. □

Ignoring user types thus reduces aggregate outcomes achieved and hurts hard-to-help users. It is necessary to pay a bonus for outcomes achieved by harder-to-help users in order to maximize aggregate outcomes. Figure 3 illustrates a suboptimal pure outcome-based incentive scheme and an optimal type-specific outcome-based scheme.

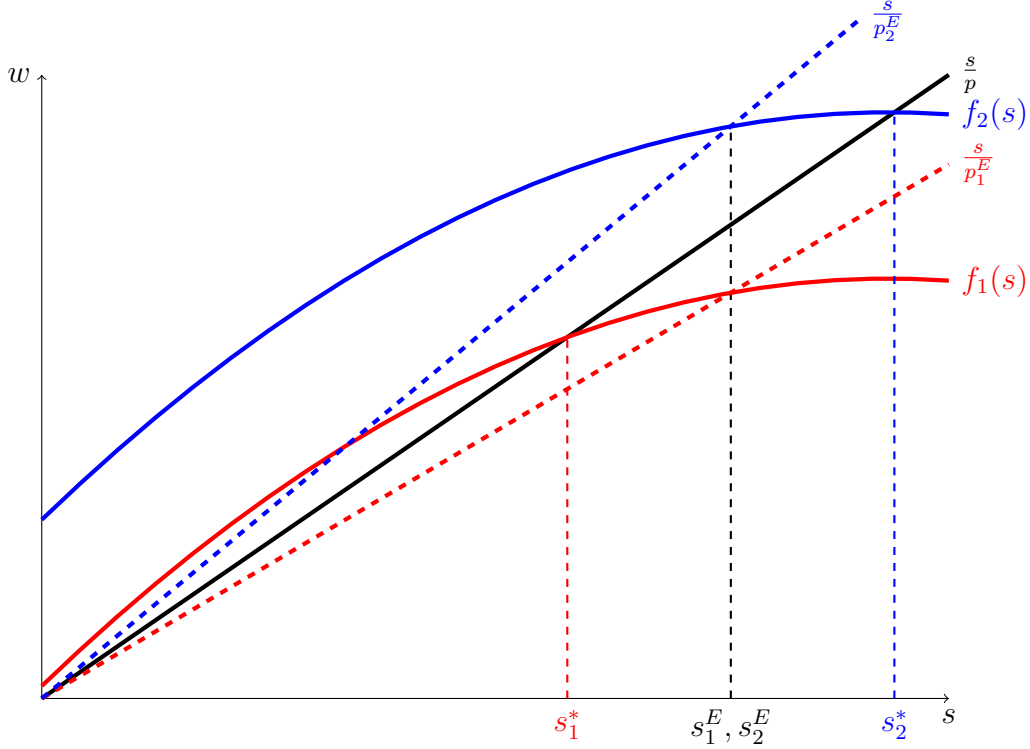


Figure 3: Outcome-Based Incentive Schemes

A pure outcome-based incentive scheme leads to overspending on the easier-to-help type relative to the harder-to-help type. Since the marginal productivity of the two types are equal for any given level of service quality, the funder optimally constructs prices such that service quality is equalized across both types.

The next incentive scheme I consider is sometimes called a value-added or a risk-adjusted reward system. Here, the performance of a given user is evaluated relative to the performance of all users of the same type, so that the better a user performs relative his peers, the higher the reward for the service provider. This scheme has been used in report cards for surgeons, where a surgeon's patient mortality rate is compared to a predicted mortality rate given the characteristics of his patients. It is also often used to compensate teachers for the performance of their students, where the outcome is defined as test score improvement.

Definition 6. *A value-added incentive scheme has the property that*

$$p(t, w; \{t_i, w_i\}_{i \notin \Psi_j}) = x_0 + h \left(w - \frac{\sum_{i \in I_t^{-j}} w_i}{I_t^{-j}} \right)$$

where I_t^{-j} denotes the set and number of users in I_t not helped by provider j , and where

$h'(\cdot) > 0$.

Note that in practice the average performance for a given type might be calculated without excluding the user in question; however, the price-taking assumption is likely innocuous as user pools are typically quite large.

Proposition 5. *A value-added incentive scheme is suboptimal. If e_i is identically distributed over types, or if h is linear, then a value-added incentive scheme provides identical service quality for all users.*

Proof. See Appendix. □

While identically distributed e_i over types is plausible in the context of student achievement, it is impossible in the case of binary outcomes such as mortality rates for patients. Report cards, however, report risk-adjusted mortality rates that are based on linear h . Given either of these conditions, value-added schemes thus result in an equal allocation of resources over all users, which will in general be suboptimal. A modified value-added scheme that was optimal would require extra reward for providers who serve users with higher optimal service quality levels. In terms of report cards for surgeons, the information provided in report cards is correct and provides valuable information to patients about the service quality of surgeons; thus, a more efficient scheme might offer financial incentives to compensate for higher effort patients. Figure 4 illustrates a value-added scheme.

The final scheme I will consider is pay-for-percentile, which pays service providers on the basis of the rank of users within a specified peer group. Here I will consider peer groups for a user to be those of the same type. This scheme was proposed by Barlevy and Neal (2012) in order to avoid the need to attach cardinal meaning to test scores, instead relying solely on the assumption that test scores preserve the ordinal ranking of human capital formation. Because binary outcomes would lead all users to achieve either the highest or lowest percentile, this scheme is unlikely to be used in such settings.

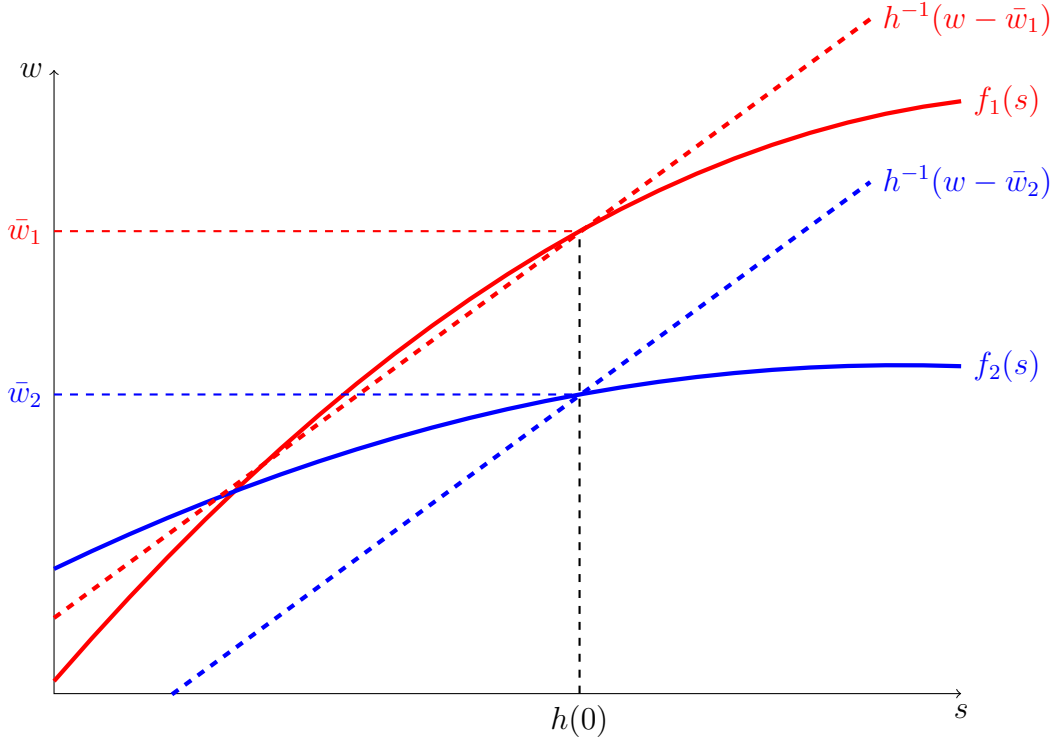


Figure 4: Value-Added Incentive Scheme

A value-added incentive scheme leads to equal service quality for all types, regardless of the production functions.

Definition 7. *A pay-for-percentile incentive mechanism has the property that*

$$p(t, w; \{t_i, w_i\}_{i \notin \Psi_j}) = x_0 + x_1 F_{I_t^{-j}}(w)$$

where $F_{I_t^{-j}}$ denotes the cumulative distribution of outcomes achieved by users of type t except those with provider j , and where $x_1 > 0$.

Importantly, notice that the fixed payment x_0 does not depend on the user's type. Barlevy and Neal (2012) acknowledge that in order to minimize the cost of implementing this mechanism, x_0 should vary over type so that each user's payoff is exactly equal to his outside option when choosing optimal quality. However, without competition for users, this scheme with constant x_0 still elicits the optimal quality levels given some lower budget—that is, while money is wasted, the distribution of quality over user types is not distorted. The less

efficient version of the scheme considered here may be more likely to be adopted in practice, in any case.¹⁰

Proposition 6. *A pay-for-percentile incentive scheme is suboptimal. If e_i is identically distributed over types, then a pay-for-percentile scheme provides identical service quality for all users.*

Proof. See Appendix. □

Just as with the value-added scheme, all users receive the same service quality in equilibrium. The intuition is the same as well; if providers offered the optimal quality levels for each user type, users requiring less quality would be more valuable and providers would offer more quality for these types of users.¹¹

3 Discussion

The simple model described in this paper demonstrates that the introduction of competition for users on the basis of service quality can have important ramifications for the design of incentive schemes. Popular incentive schemes are suboptimal and can have perverse distributional consequences. Whether this insight is important and should be incorporated into future, more complex models depends on the application. In settings where competition for users is an important feature, it should be taken into account. I consider here the markets for public school teachers and physicians, discussing the extent to which service providers may compete for users to help, whether this form of competition can help explain established facts within each market, and how competition for users should inform the design of performance-based incentive schemes.

¹⁰For example, in an experiment on the effect of teacher incentives, Fryer et al. (2012) set x_0 equal to the teacher's normal salary and $x_1 = \$8,000$.

¹¹The relevance of this result for practical design of teacher incentives will depend on the extent to which teachers can compete for students, and may vary across schools.

3.1 Teachers

In the United States, the public education system employs 3.1 million teachers for 49.5 million students (Keaton 2013). Research has shown that teachers are not only tremendously important for the educational outcomes of their students, but that they vary significantly in quality as well (e.g. Hanushek and Rivkin 2012). Such research has motivated policy reforms that attempt to decrease the disparities in educational achievement across students, as well as improve aggregate achievement. A central strategy has been the use of performance-based incentives for teachers and schools; however, these schemes have not addressed the possibility that teachers compete for students on the basis of endogenous quality, i.e., effort.

Within schools, competition among teachers for students on the basis of effort may be facilitated by principals, as they typically have large discretion in matching students with teachers (see, for example, Monk (1987)). Thus, at least in theory, a principal could reward high quality teachers with better students. Indeed, Kalogrides, Loeb and Beteille (2013) find that within a given grade level at a given school, teachers who are more experienced, have degrees from more competitive colleges, and who hold more leadership positions are assigned higher-achieving students than their peers.¹² While many schools may not offer financial incentives to teachers for the performance of their students, teachers may obtain greater utility gains from the opportunity to teach higher achieving students.

The possibility that teachers compete for students on the basis of effort is consistent with education research. Ballou and Podgursky (1995) analyze a survey of teachers that asks how many after-school hours they worked and whether they gave homework for each class they taught during the previous week. They find that, even when controlling for characteristics of the schools and teachers themselves, teachers spend more hours and are more likely to assign homework in classes with higher-achieving students. This is true when comparing variation in effort across teachers as well as for different classes taught by the same teacher, suggesting

¹²An alternative explanation for principals allocating high-achieving students to high quality teachers might be that teacher quality is marginally more effective for high-achieving students. This possibility is considered later within this section.

that the relationship is not purely due to exogenously hard-working teachers sorting with high-achieving students. Meanwhile, Clotfelter, Ladd and Vigdor (2009) find that teacher absences are more likely among teachers in schools where more students are enrolled in free lunch programs and where more students are nonwhite.

An alternative explanation for greater exertion of effort among teachers for high-achieving students is that outcomes are more responsive to effort for these students. Recent studies evaluating policy reforms and others based on field experiments that incentivize additional effort by teachers can help distinguish between these two explanations. To the extent that gains for all students are rewarded similarly, greater gains made by low-achieving students would suggest that they had a higher marginal return to teacher effort, providing evidence against this alternative hypothesis. Indeed, in a study of a teacher incentive program rewarding student achievement growth in North Carolina, Ladd and Lauen (2010) find that students with lower initial test scores experience more growth in test scores than students with high initial test scores. Figlio and Kenny (2007) use cross-sectional data on teacher incentive programs and student outcomes and find that students in the lower and middle third of the achievement distribution experience larger increases in performance than students in the upper third. Lavy (2009) finds that a teacher incentive program in Israel, which paid individual teachers bonuses based on relative performance, led to larger increases in achievement for lower-achieving students.¹³ Evidence from field experiments is mixed, with some finding greater gains at the high end of the distribution (Duflo, Hanna and Ryan 2012), and others finding no difference across the achievement distribution (Muralidharan and Sundararaman 2011; Fryer et al. 2012; Fryer 2013).

One last piece of evidence, which is unexplained in the education literature, but which is consistent with teacher competition for students on the basis of effort, is that while there is large estimated variation in teacher quality, very little of this variation is explained by

¹³However, because one component of the competition was percent of students who passed the exam, teachers would optimally expend more effort on students in danger of failing even if effort was not as effective in raising their achievement.

observed characteristics of teachers such as degrees, certifications completed, or test scores on certification exams. One explanation may be that the unexplained variation is partly due to nonrandom sorting of students across teachers of varying quality (Rothstein 2010). An alternative explanation is that teachers are actually similarly capable of high quality instruction, with the variation in observed quality due to equilibrium differences in their supply of effort. The distinction is important. If teachers have heterogeneous but fixed quality differences, then mechanisms that re-sort teachers can be successful in reducing achievement gaps. But if teachers have similar fixed quality levels, then attempting to re-sort teachers may cause them to endogenously change their quality (i.e., effort) in response to their new students.

To the extent that teachers compete for students based on effort, performance-based incentive schemes should be introduced carefully. Teachers can be offered bonuses for teaching students for whom teacher effort is marginally more effective, as well as for lower-achieving students whose success may be weighted more heavily by policymakers. This would increase the competitive pressure on teachers to exert more effort and thus maintain the opportunity to teach these higher profit students. There is some evidence that such a strategy could be successful. Clotfelter et al. (2008) find that annual bonuses of \$1,800 to North Carolina teachers working in schools with high poverty or low test scores reduced turnover rates at these schools by 17%. Whether teacher effort expands as well should be explored in future research.

3.2 Physicians

Annual healthcare expenditures in the United States total \$2.7 trillion, including close to \$1 trillion spent on government-sponsored Medicare and Medicaid programs alone (Centers for Medicare and Medicaid Services 2013). Effective care per dollar spent, however, varies tremendously across the country, with health outcomes no better in areas that spend significantly more money on healthcare (Fisher et al. 2003*a*). Moreover, patient satisfaction

and physician expressed quality of care are lower in high spending areas (Fisher et al. 2003*b*; Sirovich et al. 2006). Existing explanations focus on exogenous differences in practice styles and quality among physicians, with some more likely than others to provide unnecessary treatment (Grytten and Sorensen 2003; Cutler et al. 2013). An alternative explanation is that physicians substitute unnecessary tests and specialty visits for high effort levels with less desirable patients. In a survey of Medicare patients, Lee and Kasper (1998) find that patients who are black, older, less educated, poorer, and of lower health rate their physicians as worse in terms of “information giving” and “interpersonal manner.” If physicians tend to prefer not to treat these demographics of patients, this evidence is consistent with endogenous effort exertion by physicians.

To the extent that endogenous effort is an important factor in patient matching with physicians, incentive schemes that reward physicians for patient outcomes should account for it. Introducing performance-based incentives for patient outcomes could be helpful, but bonuses should be offered for serving patients for whom physician effort is marginally more effective.

4 Conclusion

With increasing demand for accountability for social service providers, performance-based incentives have become more common. The simple model in this paper shows that if providers compete for users to serve by adjusting service quality, then popular incentive schemes will be inefficient and may have poor distributional properties. Optimal incentive schemes must account for competition for users, and competition can even replace the need to reward performance. The extent to which service providers actually compete for users to help on the basis of service quality, and thus whether incentive schemes used in practice should account for competition, depends on the context. There is some evidence that this form of competition can be important in the markets for teachers and physicians. Future research

should further explore the role of competition in these and other social service markets, and more nuanced models that more realistically account for other features of individual settings should be taken into account in developing performance-based incentive schemes.

References

- Baker, George P.** 1992. “Incentive Contracts and Performance Measurement.” *Journal of Political Economy*, 100(3): 598–614.
- Ballou, Dale, and Michael Podgursky.** 1995. “Education Policy and Teacher Effort.” *Industrial Relations*, 34(1): 21–39.
- Barlevy, Gadi, and Derek Neal.** 2012. “Pay for Percentile.” *American Economic Review*, 102(5): 1805–1831.
- Centers for Medicare and Medicaid Services.** 2013. “National Health Expenditure Accounts.”
- Clotfelter, Charles, Elizabeth Glennie, Helen Ladd, and Jacob Vigdor.** 2008. “Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina.” *Journal of Public Economics*, 92: 1352–1370.
- Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor.** 2009. “Are Teacher Absences Worth Worrying About in the U.S.?” *Urban Institute Working Paper*.
- Cutler, David, Jonathan Skinner, Ariel Dora Stern, and David Wennberg.** 2013. “Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending.” *NBER Working Paper No. 19320*.
- Dranove, David, Daniel Kessler, Mark McClellan, and Mark Satterthwaite.** 2003. “Is More Information Better? The Effects of “Report Cards” on Health Care Providers.” *Journal of Political Economy*, 111(3): 555–588.
- Duflo, Esther, Rema Hanna, and Stephen Ryan.** 2012. “Incentives Work: Getting Teachers to Come to School.” *American Economic Review*, 102(4): 1241–1278.
- Ellis, Randall P.** 1998. “Creaming, Skimping and Dumping: Provider Competition on the Intensive and Extensive Margins.” *Journal of Health Economics*, 17: 537–555.
- Feldstein, Martin S.** 1971. “Hospital Cost Inflation: A Study of Nonprofit Price Dynamics.” *American Economic Review*, 61(5): 853–872.
- Figlio, David N., and Lawrence W. Kenny.** 2007. “Individual Teacher Incentives and Student Performance.” *Journal of Public Economics*, 91: 901–914.
- Fisher, Elliot S., David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F.L. Lucas, and Etoile L. Pinder.** 2003*a*. “The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care.” *Annals of Internal Medicine*, 138(4): 273–287.
- Fisher, Elliot S., David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F.L. Lucas, and Etoile L. Pinder.** 2003*b*. “The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care.” *Annals of Internal Medicine*, 138(4): 288–298.

- Fryer, Roland G.** 2013. "Teacher Incentives and Student Achievement: Evidence from New York City Public Schools." *Journal of Labor Economics*, 31(2): 373–407.
- Fryer, Roland G., Steven D. Levitt, John List, and Sally Sadoff.** 2012. "Enhancing the Efficacy of Teacher Incentives through Loss Aversion: A Field Experiment." *NBER Working Paper*.
- Grytten, Jostein, and Rune Sorensen.** 2003. "Practice Variation and Physician-Specific Effects." *Journal of Health Economics*, 22: 403–418.
- Hanushek, Eric A., and Steven G. Rivkin.** 2012. "The Distribution of Teacher Quality and Implications for Policy." *Annual Review of Economics*, 4: 131–157.
- Held, Philip J., and Mark V. Pauly.** 1983. "Competition and Efficiency in the End Stage Renal Disease Program." *Journal of Health Economics*, 2(2): 95–118.
- Holmstrom, Bengt.** 1979. "Moral Hazard and Observability." *The Bell Journal of Economics*, 10(1): 74–91.
- Holmstrom, Bengt, and Paul Milgrom.** 1987. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica*, 55(2): 303–328.
- Holmstrom, Bengt, and Paul Milgrom.** 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization*, 7: 24–52.
- Joskow, Paul L.** 1980. "The Effects of Competition and Regulation on Hospital Bed Supply and the Reservation Quality of the Hospital." *The Bell Journal of Economics*, 11(2): 421–447.
- Kalogrides, Demetra, Susanna Loeb, and Tera Beteille.** 2013. "Systematic Sorting: Teacher Characteristics and Class Assignments." *Sociology of Education*, 88(2): 103–123.
- Keaton, Patrick.** 2013. "Selected Statistics from the Common Core of Data: School Year 2011-2012." National Center for Education Statistics, U.S. Department of Education NCES 2013-441, Washington, D.C.
- Kessler, Daniel P., and Mark B. McClellan.** 2000. "Is Hospital Competition Socially Wasteful?" *Quarterly Journal of Economics*, 115(2): 577–615.
- Ladd, Helen F., and Douglas L. Lauen.** 2010. "Status versus Growth: The Distributional Effects of School Accountability Policies." *Journal of Policy Analysis and Management*, 29(3): 426–450.
- Lavy, Victor.** 2009. "Performance Pay and Teachers' Effort, Productivity, and Grading Ethics." *American Economic Review*, 99(5): 1979–2011.
- Lee, Yunhwan, and Judith D. Kasper.** 1998. "Assessment of Medical Care by Elderly People: General Satisfaction and Physician Quality." *Health Services Research*, 32(6): 741–758.

- Ma, Ching-To Albert, and Thomas G. McGuire.** 1997. “Optimal Health Insurance and Provider Payment.” *American Economic Review*, 87(4): 685–704.
- Monk, David H.** 1987. “Assigning Elementary Pupils to Their Teachers.” *Elementary School Journal*, 88(2): 166–187.
- Muralidharan, Karthik, and Venkatesh Sundararaman.** 2011. “Teacher Performance Pay: Experimental Evidence from India.” *Journal of Political Economy*, 119(1): 39–77.
- National Center for Education Statistics.** 2013. “The Nation’s Report Card: A First Look: 2013 Mathematics and Reading.” Institute of Education Sciences, U.S. Department of Education NCES 2014-451, Washington, D.C.
- Neal, Derek.** 2011. “The Design of Performance Pay in Education.” *Handbook of Economics of Education*, 4.
- Neal, Derek, and Diane Whitmore Schanzenbach.** 2010. “Left Behind by Design: Proficiency Counts and Test-Based Accountability.” *Review of Economics and Statistics*, 92(2): 263–283.
- Rothstein, Jesse.** 2010. “Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement.” *Quarterly Journal of Economics*, 125(1): 175–214.
- Sirovich, Brenda E., Daniel J. Gottlieb, H. Gilbert Welch, and Elliot S. Fisher.** 2006. “Regional Variations in Health Care Intensity and Physician Perceptions of Quality of Care.” *Annals of Internal Medicine*, 144(9): 641–649.
- Stecher, Brian M.** 2002. “Consequences of Large-Scale, High-Stakes Testing on School and Classroom Practice.” In *Making Sense of Test-Based Accountability in Education*. , ed. Laura S. Hamilton, Brian M. Stecher and Stephen P. Klein, 79–100. Santa Monica, CA:Rand Corporation.
- U.S. Census Bureau.** 2013. “Public Education Finances: 2011.” U.S. Government Printing Office G11-ASPEF, Washington, D.C.

Appendix

Proof of Lemma 1: Suppose $\pi_j^* > \pi_k^*$ for some $j \neq k$. Note that $\pi_k^* \geq 0$ since provider k can always select $s_k = 0$ and make 0 payoff. Also, we must have $s_j^* \neq s_k^*$ since otherwise, the payoff earned by the providers would be equal as they are equally attractive to all users. Consider the alternative strategy for program k , $s_k^{**} = s_j^* + \epsilon$, $\epsilon > 0$. Since $s_k^{**} > s_j^*$, all users originally helped by provider j would now prefer to be helped by provider k . If we make ϵ small enough, then π_k^{**} can either be made arbitrarily close to π_j^* if the incentive scheme is continuous in the neighborhood of s_j^* , or strictly greater than π_j^* if the incentive scheme is discontinuous. Either way, it follows that $\pi_k^{**} > \pi_k^*$, contradicting inequality in equilibrium payoff. \square

Proof of Proposition 1: Suppose $\pi_j^* > 0$ for some $j \in J$, and recall that $J > T$. Note that we cannot have $s_j^* = s_k^*$ for any $j, k \in J$, since then, provider k could change its strategy to $s_k^{**} = s_k^* + \epsilon$ and increase its payoff by picking ϵ small enough, just as in the proof of Lemma 1. Thus we know that $s_j^* \neq s_k^* \forall k \in J$. But since $J > T$, there must be some type t which more than one provider is helping; otherwise some provider would be helping no users, earn 0 payoff and thus contradict Lemma 1 since we assumed $\pi_j^* > 0$ for all $j \in J$. But if two programs j and k are helping the same type of user, then $s_j^* = s_k^*$; otherwise, the provider with the lower services would have no users. However, this equality contradicts what we just showed—that $s_j^* \neq s_k^*$. From this contradiction, it follows that $\pi_j^* = 0 \forall j \in J$. \square

Proof of Corollary 1: Suppose $t_i = t_{i'}$, with $i \in \Psi_j^*$, $i' \in \Psi_k^*$. Now suppose $s_k^* > s_j^*$. Since $\pi_j^* = \pi_k^* = 0$, we must have $E[p(t_i; w_i)] = s_k^*$. By our earlier assumption, this implies provider k is willing to accept user i , and thus, $i \notin \Psi_j^*$. This contradiction proves that $s_k^* = s_j^*$. \square

Proof of Proposition 2: Any optimal incentive scheme is a solution to the problem

$$\begin{aligned} & \max_{\{P_j\}_{j \in J}} \sum_{i \in I} \lambda_{t_i} f_t(s_{t_i}) \\ & \text{s.t.} \quad \sum_{i \in I} E[p(t_i, w_i)] \leq M \\ & \quad \quad E[p(t_i, w_i)] = s_{t_i} \forall i \in I \end{aligned}$$

where the first constraint is the funder's budget constraint and the second set of constraints is a consequence of competition. Combining the constraints and solving, the first order conditions imply that $\lambda_1 f'_1(s_1^E) = \lambda_t f'_t(s_t^E)$ for all $t > 1$ and $\sum_{i \in I} s_{t_i}^E = M$. If we let $E[p(t_i, f_t(s_{t_i}^E))] = s_{t_i}^E$ and $E[p(t, f_t(s))] < s$ for all $s > s_{t_i}^E$, we know from Proposition 1 that $s^* = s^E$. This shows the “if” part. The “only if” part can be seen by noting that either $E[p(t_i, f_t(s_{t_i}^E))] > s_{t_i}^E$ or $E[p(t, f_t(s))] \geq s$ for some $s > s_{t_i}^E$ implies $s^* > s_{t_i}^E$. \square

Proof of Proposition 3: While there are many examples of production functions which will show level-based incentive schemes to be suboptimal, we will consider the case where $f_t(s) = f_{t-1}(s) + \alpha_t$ where $\alpha_t > 0$, for all $t \in T$, and with $e_i = 0$. Further suppose there exists some \bar{t} such that $f_{\bar{t}-1}(x_2) < \bar{w}$ and $f_{\bar{t}}(x_2) > \bar{w}$. This implies that $s_{\bar{t}-1}^* \leq s_{\bar{t}}^*$ for all $t \in T$ and $s_{\bar{t}-1}^* < s_{\bar{t}}^*$. Optimal schemes require that $s_t^* = s_r^*$ for all $t \in T$, and so level-based incentive schemes are suboptimal and underserve hard-to-help users. \square

Proof of Proposition 4: As long as $E[h(t_i, f(s_{t_i}^E) + e_i)] = s_{t_i}^E$ and $E[h(t_i, f(s_{t_i}^E) + e_i)] < s$ for all $s > s_{t_i}^E$, then by Proposition 2 this outcome-based scheme will be optimal. Now we show that a pure outcome-based scheme is suboptimal and underserves hard-to-help users. Suppose, for example, that $f_t(s) = f_{t-1}(s) + \alpha_t$ where $\alpha_t > 0$, for all $t \in T$, and $e_i = 0$. Since $h(t_i, w_i) = h(w_i)$, it follows that $s_t^* = h(f_t(s_t^*)) = h(f_{t-1}(s_{t-1}^*) + \alpha_t) > h(f_{t-1}(s_{t-1}^*)) = s_{t-1}^*$. Therefore, a pure-outcome based incentive scheme is suboptimal and underserves hard-to-help users. \square

Proof of Proposition 5: Suppose an equilibrium exists.¹⁴ Note that from Corollary 1 we know that all users of a given type receive the same service quality level, which implies that $s_t^* = x_0 + E[h(e_i)]$ for all $t \in T$. If e_i is identically distributed over types, then $s_t^* = s^*$ for all types $t \in T$. If h is linear, then $s_t^* = x_0 + E[\alpha e_i] = x_0 + \alpha E[e_i] = x_0$, and so all users receive the same service quality. This is optimal only in the special case where $s_r^E = s_t^E$ for all $r, t \in T$. \square

Proof of Proposition 6: Suppose an equilibrium exists.¹⁵ Now, note that from Corollary 1 we know that all users of a given type receive the same service quality level, which implies that $s_t^* = x_0 + x_1 E[F_{T-t}(e_i)] = x_0 + \frac{1}{2}x_1$ for all $t \in T$, using the assumption that e_i is identically distributed over types. Thus, all users receive the same service quality. This is optimal only in the special case where $s_r^E = s_t^E$ for all $r, t \in T$. \square

¹⁴Existence requires that $h(w)$ does not grow too quickly. Since all users of a given type must be treated equally in equilibrium, if h grows too quickly, then providers will wish to deviate from this strategy.

¹⁵Existence will depend on the magnitude of $\text{Var}(e_i)$. To see why, consider the case where $\text{Var}(e_i) = 0$. Then, if all providers are treating a given type the same (as must be true in equilibrium), a small deviation in service quality would move the user from the 50th percentile to the highest percentile, and thus, this cannot be an equilibrium.