Choice and Competition in Public Service Provision*

Timothy Besley
LSE and CIFAR

James M. Malcomson
University of Oxford

August 16, 2016

Abstract

In spite of a range of policy initiatives in sectors such as education, health care and legal services, whether choice and competition is valuable remains contested territory. This paper studies the impact of choice and competition on different dimensions of quality, examining the role of not-for-profit providers. We explore two main factors which determine whether an alternative provider enters the market: cost efficiency and the preferences of an incumbent not-for-profit provider (paternalism). The framework developed can incorporate standard concerns about the downside of choice and competition when consumer choice is defective (an internality) or choice imposes costs on those who do not switch (an externality). The paper considers optimal funding levels for incumbents and entrants showing when the "voucher" provided for consumers to move to the incumbent should be more or less generous than the funding for consumers who remain with the incumbent. Finally, the model also offers an insight into why initiatives are frequently opposed by incumbent providers even if the latter have not-for-profit objectives.

Keywords: Choice, Competition, Public Service, Not-for-profit

JEL codes: H11, H44, L21, L31.

*We are grateful to Kurt Brekke, Martin Chalkley, Steve Coate, Marty Gaynor, Julian Le Grand, Carol Propper, Luigi Siciliani and participants in the Economic Theory Workshop at Nuffield College, Oxford for helpful comments and discussion.

1 Introduction

Across the world, democratically-elected governments face voter pressure to provide high quality public services at reasonable cost. Many public services have traditionally been provided by state-funded local monopolies. One policy option is to open these up to competition, based on the widespread view that competition provides the services consumers want at the lowest cost. But this policy is controversial, at least in part because many public services are provided by not-for-profit providers, whereas potential entrants may be more interested in profits and less concerned with the quality of services. This is potentially important where, as in many public services, there are difficulties in ensuring quality, externalities, equity concerns and minimum service obligations that not-for-profit providers are, in the view of many, better for overcoming. But not-for-profit providers may use a monopoly position to further their own provider interests, not those of consumers. This paper develops a framework to investigate the implications of competition in such services that allows for the potential advantages of not-for-profit provision.

A classic debate along these lines concerns the use of education vouchers to enhance competition. Some influential commentators have argued that allowing competitive for-profit provision can be welfare enhancing (see, for example, Tooley and Dixon (2005)). Even without encouraging for-profit provision, countries such as Sweden have encouraged entry of "free schools" financed with publicly-funded vouchers (see, for example, Böhlmark and Lindahl (2015)). In similar vein, health care systems that have traditionally relied on monopoly public provision are considering allowing for-profit providers to compete to provide services. Another relevant example is the provision of legal services for low-income individuals accused of a crime in order to ensure them a fair trial. Here a key question is whether standard for-profit law firms should be allowed to provide these services.

Our model of choice and competition in public services has four key features. First, there are "mobile" consumers who are willing to exercise a choice to switch providers. Second, potential service providers differ in their costs, with some having an intrinsic cost advantage. Third, there are dimensions of quality that consumers are not in a position to assess before choosing a provider, which we refer to as *unobservable* quality. Fourth, not-for-profit providers differ from for-profit providers in their concerns for quality but those concerns need not be aligned with consumer preferences (possibly because of provider paternalism). Under not-for-profit provision we include state provision that, even if not by independent not-for-profit firms, have decisions decentralized to, for example, school principals and hospital administrators.

We begin with the simplest model that illustrates our main insights, generalizing it in an appendix to show that the key logic applies more widely. We start with an incumbent monopoly provider who receives a fixed payment per consumer of its service. Examples of such payments are the fixed payment per patient for a given treatment under Medicare in the US (payment by *diagnosis related group*) and the British National Health Service (NHS). In this setting, consumers are better off with a not-for-profit provider no matter how much more efficient an alternative for-profit provider might be because the not-for-profit provider provides more than minimal, even if far from optimal, quality. In contrast, a for-profit provider's only interest is to cut the cost of provision and thus provide minimal quality, both observable and unobservable. We then derive conditions under which entry, if permitted, will occur. If both incumbent and entrant are for-profit, whether entry occurs depends only on their relative efficiencies and the proportion of consumers who are mobile. Unobservable quality remains at a minimal level but those consumers who switch to the entrant benefit from an increase in observable quality and are better off as a result. If the incumbent is not-for-profit, the extent to which its preferences are aligned with those of consumers also affects whether entry occurs. A for-profit entrant then requires a greater efficiency advantage for entry to be worthwhile because the not-for-profit incumbent has an effective cost advantage from providing unobservable quality that consumers value at low marginal cost. But when entry occurs, consumers who switch are better off because of higher observable quality, albeit at the expense of lower unobservable quality. Importantly, all consumers are better off than if the not-for-profit incumbent was replaced by an equally productive for-profit, with the policy implication that it is beneficial to ensure at least one not-for-profit provider remains in business. In contrast, when a not-for-profit incumbent faces a not-for-profit entrant, their paternalism can induce entry even if they are equally efficient because the entrant offers consumers who switch a balance between observable and unobservable qualities that is closer to their preferences than a monopoly not-for-profit incumbent would have done. Then entry is a discipline on the "decision rents" earned by monopoly incumbents, creating an additional benefit to consumers from allowing choice and competition. Two key messages are that not-for-profit provision is valuable even when there is competition and that, for assessing the impact of competition on quality of service, it is important to distinguish quality dimensions that are readily discernible by consumers from those that are not.

Competition enables consumers who switch to achieve higher utility for themselves *as perceived by them.* But there are concerns about the value of consumer sovereignty in some public services. Consumers who switch providers may impose an externality on those who do not switch because of, for example, peer effects in education. They may also not fully appreciate the value of some dimensions of quality — for example, the benefits of a small reduction in wound infection rates versus additional convenience for relatives to visit them in hospital. We discuss formally how these can undermine the case for choice and competition. In the second of these, a monopoly provider's paternalism can be beneficial; competition reduces the scope for the exercise of such paternalism.

With traditional voucher systems for education, health and legal services, an entrant who attracts customers receives the same fee per customer as the incumbent. But is creating such a "level playing field" optimal? Differential payment to an entrant can be an important policy tool for promoting or deterring choice and competition. We formulate the optimal funding problem and show that a level playing field is not typically optimal. We then identify the factors that shape the optimal voucher for customers who stay with the incumbent, including the shape of the distribution of cost efficiency among potential entrants and the size of gains to consumers who switch. We also show that the optimal funding level for the incumbent should depend on whether competition is permitted.

Finally, the model provides an insight into the political economy of public service reform and why opposition to competition should be expected from incumbents whose rents (whether earned as profit or having control over decisions) are threatened. A case in point is the opposition to charter schools by teachers' unions. It is important that, even if there are caveats to welfare consequences due to internalities and externalities, such insiders are not the only decisive party in determining competition policy even when they have a *bona fide* not-for-profit objective.

The rest of the paper is organized as follows. In the next section, we discuss related literature. Section 3 introduces the core modeling framework. It also sets up the monopoly benchmark and motivates the role for not-for-profit provision in that framework. Section 4 allows entry and studies choice and competition. Section 5 develops implications for different provider objectives. In section 6, we discuss two standard caveats to the argument that permitting entry raises consumer welfare, internalities and externalities. Section 7 develops an analysis of optimal funding including the optimal voucher that should be offered to consumers who move to an entrant. Section 8 discusses provider interests and why our framework predicts that they will tend to go against allowing choice and competition in public services even when consumers benefit. Section 9 concludes. Appendix A contains proofs of propositions. Appendix B shows that the main results are robust to allowing for a more general objective function for not-for-profit firms, a continuous distribution of switching costs/benefits for consumers and more than two quality dimensions.

2 Related Literature

The paper is related to the large literature on the merits of not-for-profit service provision. We build on two established traditions in the study of not-for-profits. From Newhouse (1970), we use the idea that not-for-profit providers have a bias towards quality relative to for-profit providers and, following Hansmann (1980), we acknowledge the importance of the non-contractibility of quality in understanding why firms choose not-for-profit status. The latter is used in Glaeser and Shleifer (2001) and lies behind the core trade-offs uncovered in Hart et al (1997). The key point is that there is a potential cost-quality trade-off. One way to mitigate the trade-off is to employ motivated agents who care directly about quality, as in Besley and Ghatak (2001). Ghatak and Mueller (2011) show that the selection of motivated workers is an advantage to a not-for-profit firm.

Here we study the impact of competition when not-for-profit providers have the characteristics emphasized by Newhouse (1970) and Hansmann (1980). The role of competition in public service provision has been discussed in Le Grand (2007). Hoxby (1999) has discussed some formal models of how competition can matter. Lakdawalla and Philipson (2006) also discusses competition with a not-for-profit provider. In that model, a not-for-profit differs from a for-profit only in having the quantity it provides as an argument in its objective function in addition to, and separate from, its role in generating profit. Only because charitable donations enable it to operate at a loss can it indulge its own preferences relative to a for-profit provider with the same cost function. Quality of service does not enter the model. More recently, Laine and Ma (2016) include quality of service in their model of competition between public and private firms. Their public firms, however, are assumed to maximize social surplus, which makes them very different from the not-for-profit providers in Newhouse (1970) that have their own self interests.

Brekke et al (2011) and Brekke et al (2012) study the effect of competition on quality with not-for-profit providers modeled as caring about consumer benefits in addition to profit. They focus on strategic interaction where providers take simultaneous decisions (although there may be several stages to the game where quality, price and/or location are chosen) and show that the impact of competition on equilibrium quality requires a subtle understanding of features of the market and provider behavior. In an application to health care, Brekke et al (2014) have considered how patient mobility affects provision of health care when governments make quality investment decisions to maximize welfare and study the important question of how this depends on transfer payments when patients shop around. In these models, however, quality has a single dimension observable by consumers, so there is not the underlying rationale for not-for-profit providers emphasized in Hansmann (1980).

The analysis of competition and entry in education is extensive. In its early incarnation, the focus was on competition between jurisdictions with population mobility. However, in recent years interest has been fuelled in large measure by the US charter school experiment allowing entry of schools to compete against public providers. The latter has been taken up in a range of countries including Sweden and the UK. There is now a large theoretical and empirical literature on the role of competition in improving the performance of schools. From the theoretical side, there are contributions by Barseghyan et al (2014), Epple and Romano (1998) and McMillan (2005). Empirical studies of the impact of school competition include Card et al (2010), Hoxby (2003), Lavy (2008) and Gibbons et al (2008). However, as yet there is no canonical theoretical approach to entry in competition with public providers which models how this works and which takes into account of the possibility of strategic interaction between providers.

The paper is also related to the large literature on school vouchers (see Ladd (2002) and Neil (2002) for reviews) following the early advocacy of the idea by Friedman (1962). Standard models, such as Nechyba (2000), look at the possibility that a citizen can carry their public funding to another provider. Böhlmark and Lindahl (2015) evaluate Sweden's school voucher system arguing that increased school competition enhanced standards. The debate about the value of voucher systems has typically centred on changes in quality and/or the gains from competition. Here we raise an additional issue — whether vouchers should be more or less generous than the capitation fee given to incumbents — and show that, because quality may not be optimal in the first place, there may be a case for either more or less generous funding of entrants relative to incumbents. We also investigate what optimal voucher design would look like when consumers do not fully appreciate the implications of all dimensions of quality and/or there is an externality due to exit from the incumbent, as in the standard peer effects model.

How to ensure service quality is also a major focus of the literature on health care, with significant implications for public provision of health services, see Chalkley and Malcomson (2000). The growing literature on the effects of competition on quality in provision of health services is reviewed in Gaynor et al (2015). The models of quality determination by providers reviewed there focus on a single quality dimension observed by customers, so again there is not the underlying rationale for not-for-profit providers emphasized in Hansmann (1980), and monopolistic competition, in which there is no strategic interaction between providers. The absence of strategic interaction seems most appropriate when there is a large number of competitors, none of which impact more on one rival than on another. If there are few competitors or location is important, a model with strategic interactions seems more appropriate. In our setting which begins with a status quo of a monopoly state-funded incumbent, modeling strategic interaction is unavoidable. The health literature also provides evidence that significant numbers of patients really do switch providers in response to competition, see Chandra et al (2016) for the US and Gaynor et al (2012) for the UK.

3 The Model

Set-up The basic model considers provision of a public service for which there is a single incumbent provider, denoted by I, and a single potential entrant, denoted by E. The service has two dimensions of quality in amounts $q, Q \geq 0$ but providers can commit to only one of these, q, before consumers choose which provider to use. To capture this, we suppose that providers choose Q only *after* consumers have chosen where to consume. (An equally good alternative would be that consumers are unable to observe Q before experiencing it.) We refer to Q as *unobservable* quality. That dimensions of quality are unobserved motivates the value of not-for-profit provision in this setting since for-profit firms have no incentive to provide such quality.

Revenue per customer is fixed at $p_i > 0$, for $i \in \{I, E\}$ and is funded from taxation. In the case of entry in education, it can be thought of as a voucher which a consumer can use to spend the per capita cost of provision with an entrant instead of remaining with the incumbent. In the case of health, it corresponds to the payment under Medicare per patient in a given *diagnosis related group* and to the *payment by results* for specific treatments under the British NHS. Thus, the model is one of decentralized service provision with centralized finance.[1]

Unobserved dimensions of quality are a characteristic feature of many public services. While a parent may be able to see what is on the curriculum that they choose for their child, whether the teachers are enthusiastic and/or knowledgeable in the subject that they teach cannot be observed ex ante. Similarly, a patient choosing a hospital may observe the level of cleanliness and even the track-record of the surgeons but will find it difficult to assess what efforts are put into patient aftercare and "softer" aspects of care such as bedside manner. Finally, someone who receives legal counsel funded by the state can see what the qualifications of the lawyer are but not how much time is set aside for such activities and whether it is simply viewed as a chore by those assigned to such work.

There is a continuum of consumers of the public service, each of whom consumes at most one unit and from that receives utility $Q + q$. In the basic model, a proportion $1 - \gamma$ of consumers are rigid in the sense of always choosing the incumbent provider as long as it offers utility of at least zero, whereas the remainder are flexible, choosing whichever of the available providers' quality bundles yields the higher utility.

Allowing for rigid consumers is a non-standard, but realistic, feature of the set-up. Many markets for public services opened up to competition have seen quite limited take-up. While this could be interpreted as consumers being content with the service they are provided, it is also interpreted as inertia. These could be real costs as when a patient must travel to receive medical treatment. However, costs could also be psychological, with consumers simply unwilling to explore alternatives even when it is in their interest to do so. Having two groups of consumers in our core model is clearly a simplification; we generalize the model to consumers with a continuum of switching costs and multiple dimensions of quality in Appendix B.

The cost of providing a unit of the service is $[c(Q) + c(q)] / \theta_i$, where $c(\cdot)$ is strictly increasing and strictly convex with $c(0) = c'(0) = 0$ and $\theta_i \in [\underline{\theta}, \bar{\theta}]$ for $i \in \{I, E\}$

[1]The model could straightforwardly be extended to allow for a regulated user fee.

an efficiency parameter that can differ between providers.[2] Making the cost function additive and identical for the two kinds of quality is a simplification that is relaxed in the generalization in Appendix B.

We assume that θ_I is known to policy makers and to any potential entrant. The entrant's cost θ_E is drawn from a distribution $G(\theta_E)$ with support $[\underline{\theta}, \bar{\theta}]$ and continuous density $g(\theta_E)$. This distribution captures uncertainty about the costs of potential entrants. We let $x_i \in [0, 1]$ denote provider i's market share.

For-Profit Provision A for-profit provider's objective is

$$\{p_i - [c(Q) + c(q)]/\theta_i\} x_i ,$$

i.e. the revenue per customer served less the cost of provision, multiplied by its market share. Consider a consumer with an outside option of U (which could include not consuming at all in which case $U = 0$). With Q set only after consumers choose a provider, a for-profit provider will set $Q = 0$, i.e. it always provides the lowest level of unobserved quality. Then observed quality and utility are the same thing, i.e.

$$q = U. \tag{1}$$

Profit per consumer at this utility level is therefore

$$v_i^{FP}(U, \theta_i, p_i) = p_i - \frac{c(U)}{\theta_i}. \tag{2}$$

For the analysis that follows, it is useful to use (2) to define $\tilde{U}_i^{FP}(\theta_i, p_i)$ by

$$p_i = \frac{c\left(\tilde{U}_i^{FP}(\theta_i, p_i)\right)}{\theta_i} \tag{3}$$

as the highest utility a for-profit provider with efficiency parameter θ_i is able to deliver without making a loss. This plays a key role in the entry analysis below.

Not-for-profit Provision We suppose that a not-for-profit provider cares about quality with its objective being a weighted sum of consumer and provider preferences, i.e.

$$[\lambda(\beta Q + q) + (1-\lambda)(Q+q)]x_i = (\alpha Q + q)\, x_i, \tag{4}$$

where $\alpha = \lambda\beta + (1-\lambda)$. The parameter λ is the weight a not-for-profit provider puts on its own preferences relative to those of consumers and β reflects the weight it puts on unobservable, relative to observable, quality which may differ from that of

[2] Any fixed cost per consumer that is independent of quality can be deducted from revenue in specifying p_i.

consumers. For $\lambda = 0$, $\alpha = 1$ and the provider is fully benevolent in the sense of maximizing consumer utility.

We focus throughout on the case where $\alpha > 1$ which is implied by setting $\beta > 1$. This allows us to capture the spirit of the classic contributions to the study of not-for-profit providers such as Newhouse (1970) and Hansmann (1980) where the provision of (unobserved) quality is the *sine qua non* of not-for-profit status. With $\beta > 1$, an increase in λ (and thus α) leads to a larger divergence between the provider's and the consumers' objectives. The model captures a key delegation problem which typifies public service provision where provider interests (for better or worse) play a key role in the way that services are provided. As we shall see below, competition can reduce the power of provider interests.[3]

A not-for-profit provider must cover its costs, which gives the breakeven constraint

$$\{p_i - [c(Q) + c(q)] / \theta_i\} x_i \geq 0. \tag{5}$$

This rules out the possibility that it receives donations to support its activities over and above the publicly funded capitation fee.[4] Since it cares directly about both kinds of quality, it chooses values of $\{Q, q\}$ to maximize (4) subject to the breakeven constraint (5) and to offering utility to attract consumers. The first-order conditions for its quality choices for given market share $x_i > 0$ are

$$\alpha\theta_i - \mu c'(Q_i) = 0 \qquad \text{and} \qquad \theta_i - \mu c'(q_i) = 0, \tag{6}$$

where μ is a Lagrange multiplier on the breakeven constraint (5). Denote the solution by $\{Q^*(\alpha, \theta_i, p_i), q^*(\alpha, \theta_i, p_i)\}$ and let

$$U^*(\alpha, \theta_i, p_i) = Q^*(\alpha, \theta_i, p_i) + q^*(\alpha, \theta_i, p_i) \tag{7}$$

denote the resulting level of consumer utility.[5] When the outside option U satisfies $U \leq U^*(\alpha, \theta_i, p_i)$, as is the case when there is no entry so $U = 0$, this is the optimal solution. Otherwise, the optimal solution is fully determined by the solution with $Q \geq q$ to the binding utility and breakeven constraints and, hence, the following pair of conditions

$$\hat{Q}^*(U, \theta_i, p_i) = U - \hat{q}^*(U, \theta_i, p_i) \tag{8}$$

$$p_i = \frac{c(U - \hat{q}^*(U, \theta_i, p_i)) + c(\hat{q}^*(U, \theta_i, p_i))}{\theta_i}. \tag{9}$$

---

[3]The model can be extended to not-for-profits with some pure managerial slack.

[4]It is straightforward to allow this possibility which is considered by Lakdawalla and Philipson (2006).

[5]Both $Q^*(\alpha, \theta_i, p_i)$ and $q^*(\alpha, \theta_i, p_i)$ are unique because $c(.)$ is strictly convex and are strictly positive because $c'(0) = 0$.

Note that $\hat{Q}^*$ is strictly positive and depends on U but is independent of α.

Analogous to what we had for a for-profit, define $\tilde{U}_i^{NP}(\theta_i, p_i)$ by

$$p_i = \frac{2c\left(\tilde{U}_i^{NP}(\theta_i, p_i)/2\right)}{\theta_i} \tag{10}$$

as the highest utility a not-for-profit provider with efficiency parameter θ_i can feasibly deliver given the breakeven constraint, i.e. where $q = Q$ as desired by consumers. For the same efficiency θ_i, it is immediate that $\tilde{U}_i^{NP}(\theta_i, p_i) > \tilde{U}_i^{FP}(\theta_i, p_i)$ because the cost function is strictly convex and the not-for-profit provider provides both types of quality. Because a not-for-profit provider's preferences ensure that it delivers positive unobservable quality, it enjoys an effective cost advantage.

A not-for-profit provider's payoff per consumer served is

$$\begin{aligned} &v_i^{NP}(U, \theta_i, p_i) \\ &= \begin{cases} \alpha Q^*(\alpha, \theta_i, p_i) + q^*(\alpha, \theta_i, p_i), & \text{if } U \in [0, U^*(\alpha, \theta_i, p_i)]; \\ \alpha \hat{Q}^*(U, \theta_i, p_i) + \hat{q}^*(U, \theta_i, p_i), & \text{if } U \in \left[U^*(\alpha, \theta_i, p_i), \tilde{U}_i^{NP}(\theta_i, p_i)\right]; \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{11}$$

It is straightforward to check that, for $U \in \left[U^*(\alpha, \theta_i, p_i), \tilde{U}_i^{NP}(\theta_i, p_i)\right]$, v_i^{NP} is a decreasing function of U and everywhere non-negative, which implies that a not-for-profit provider will always wish to be active in the market.

Monopoly Benchmark We consider as a starting point consumer utility when only one provider offers the service.[6]

Proposition 1 *With a monopoly provider, the utility it offers consumers is $u^{FP}(\theta_I, p_I) = 0$ for all (θ_I, p_I) if it is for-profit and $u^{NP}(\theta_I, p_I) = U^*(\alpha, \theta_I, p_I) > 0$ for all (θ_I, p_I) if it is not-for-profit. The utility $U^*(\alpha, \theta_I, p_I)$ is increasing in θ_I and p_I and decreasing in α.*

A monopoly for-profit provider's only interest is to minimize the cost of provision, so it offers only the lowest utility for which consumers will seek provision, normalized as zero, whatever θ_I and p_I are. As explained above, a monopoly not-for-profit provider offers consumer utility $U^*(\alpha, \theta_I, p_I)$ defined in (7) which is strictly greater than zero for all θ_I and p_I. An implication of Proposition 1 is that consumers are always better off with a not-for-profit provider, no matter how inefficient, than with a for-profit provider, no matter how efficient. Although not included formally in our model, this result carries over straightforwardly to the case in which there is a fixed cost of provision independent of the number of consumers served. If that fixed cost is sufficiently large that the market can sustain only one provider, it is always better for

[6]All proofs are in the Appenidx.

consumers that this is a not-for-profit provider. This has an direct policy implication. If a community is too small to sustain more than one school or hospital, it is better for consumers to require that the school or hospital is not-for-profit.

Another implication of Proposition 1 is that neither increasing funding (higher p_I) nor having a more efficient provider (higher θ_I) makes a difference to the quality of service supplied by a monopoly for-profit provider. In contrast, both are unambiguously better for consumers with a monopoly not-for-profit provider because they allow more of both kinds of quality to be provided. But not-for-profit provision does not maximize consumer utility for given funding because, with $\alpha > 1$, there is non-alignment between the provider's and consumers' objectives. Proposition 1 shows that consumers will actually be *worse off* if an incumbent not-for-profit either cares more about provider objectives (higher λ) and/or its bias towards quality Q is greater (higher β), either of which implies higher α. This is an important distortion that motivates a role for competition beyond achieving cost-efficiency. The rents earned by monopoly not-for-profit providers are decision rents due to their ability to determine the mix of qualities they prefer.

4 Entry

Entry serves two possible roles. First, an entrant may be more efficient (have high θ). Second, an entrant may deliver an outcome that is closer to what (flexible) consumers want.

Timing The timing is as follows:

1. Nature determines the efficiency of the potential entrant $\theta_E \in \left[\underline{\theta}, \bar{\theta}\right]$.
2. The potential entrant decides whether to enter and, if it decides to do so, chooses q_E, which is observed by consumers and the incumbent. (If the entrant anticipates the same equilibrium payoff from entering as from not entering, it chooses to enter if and only if it actually attracts some consumers.)
3. The incumbent chooses q_I, which is observed by consumers.
4. Consumers choose whether to consume and if so where with, for simplicity, indifferent flexible consumers choosing the entrant.
5. Provider I chooses Q_I and provider E, if entered, chooses Q_E.

We solve for a sub-game perfect equilibrium.

Core Entry Result For each organizational form $j \in \{FP, NP\}$, $u^j(\theta_i, p_i)$ specified in Proposition 1 for $i \in \{I, E\}$ is the utility to consumers delivered by a type j provider if not constrained by competition. When deciding whether to accommodate entry, the incumbent must decide whether to allow the entrant to serve the flexible consumers. Whether it does so depends on the difference in its payoff, whether from profits or its payoff as a not-for-profit, from serving the whole market compared to serving only the proportion $(1-\gamma)$ of the market consisting of rigid consumers. To serve the whole market, it must offer all consumers utility that matches the utility offered by the entrant. But, if it seeks to retain only the rigid consumers, it can do that by offering just $u^j(\theta_I, p_I)$ and thus receive a higher payoff per consumer served. There is thus a critical proportion of flexible consumers that makes it optimal to compete for them.

Formally, the incumbent's payoff is $(1-\gamma)\, v_I^j(u^j(\theta_I, p_I), \theta_I, p_I)$ when serving only the rigid consumers. Its payoff when serving the whole market at a utility level U determined by the entrant's offer is $v_I^j(U, \theta_I, p_I)$. The critical value of γ below which the incumbent prefers to serve only the rigid consumers is $\widehat{\gamma}^j(U, \theta_I, p_I)$ defined by

$$\widehat{\gamma}^j(U,\theta_I,p_I) = \begin{cases} 1\,, & \text{if} \quad v_I^j\,(U,\theta_I,p_I) < 0; \\ 1 - \frac{v_I^j(U,\theta_I,p_I)}{v_I^j(u^j(\theta_I,p_I),\theta_I,p_I)}\,, & \text{if} \quad 0 \le v_I^j\,(U,\theta_I,p_I) < v_I^j\left(u^j(\theta_I,p_I),\theta_I,p_I\right); \\ 0\,, & \text{if} \quad v_I^j\,(U,\theta_I,p_I) \ge v_I^j(u^j(\theta_I,p_I),\theta_I,p_I). \end{cases} \tag{12}$$

That is, if $\gamma < \widehat{\gamma}^j(U, \theta_I, p_I)$, there are too few flexible consumers for it to be worth the incumbent competing for them by offering them a payoff of U. The top and bottom cases in (12) are corner solutions where either the incumbent never finds it worthwhile to compete (top case) or always retains the flexible consumers (bottom case). As U increases, the critical value of $\widehat{\gamma}^j(U, \theta_I, p_I)$ increases and the incumbent is in a weaker position to compete. Define $\overline{U}^j(\gamma, \theta_I, p_I)$ by

$$\gamma = \widehat{\gamma}^j(\overline{U}^j(\gamma,\theta_I,p_I),\theta_I,p_I) \tag{13}$$

as the highest utility the incumbent is willing to offer to retain the flexible consumers. Note that $\overline{U}^j(\gamma, \theta_I, p_I) > u^j(\theta_I, p_I)$ because the incumbent is always willing to give up a small amount of payoff per consumer served to acquire the discrete proportion γ of flexible consumers.

The next proposition gives a necessary and sufficient condition for entry and specifies how consumers fare with and without entry. Recall that $\tilde{U}_i^j(\theta_i, p_i)$, for $i \in \{I, E\}$, is the highest utility provider type j can provide without making a loss.

Proposition 2 *Entry by type k occurs with incumbent type j, for $j, k \in \{FP, NP\}$, if and only if*

$$\gamma \le \widehat{\gamma}^j(\widetilde{U}_E^k(\theta_E,p_E),\theta_I,p_I). \tag{14}$$

If no entry occurs, payoffs for both rigid and flexible consumers are $u^j(\theta_I, p_I)$. If entry occurs, rigid consumer payoffs are $u^j(\theta_I, p_I)$ and flexible consumer payoffs are as follows:

$$\begin{array}{ll} \max\left\{\widetilde{U}_I^j(\theta_I, p_I), u^k(\theta_E, p_E)\right\}, & \text{if } \gamma \geq \widehat{\gamma}^j\left(\widetilde{U}_I^j(\theta_I, p_I), \theta_I, p_I\right); \\ \max\left\{\overline{U}^j(\gamma, \theta_I, p_I), u^k(\theta_E, p_E)\right\}, & \text{otherwise.} \end{array} \tag{15}$$

Entry strictly increases the utility of flexible consumers while leaving the utility of rigid consumers unchanged.

This result applies for all possible organizational forms and efficiency levels for the incumbent and entrant. To understand it, note that $\widetilde{U}_E^k(\theta_E, p_E)$ determines how hard the potential entrant can compete for flexible consumers since it is the highest level of utility that it can offer them and still be worth entering. The key issue is whether the proportion of flexible consumers γ is greater than $\widehat{\gamma}^j(\widetilde{U}_E^k(\theta_E, p_E), \theta_I, p_I)$. If it is, there is no entry because it is worthwhile for the incumbent to compete and retain the flexible consumers by offering them more than the highest utility the potential entrant can afford to offer. In this case, the potential entrant would be unable to capture any of the market and would not enter. If γ is below $\widehat{\gamma}^j(\widetilde{U}_E^k(\theta_E, p_E), \theta_I, p_I)$ (condition (14)), the entrant can attract the flexible consumers. But it is worth entering to do that only if it has a positive payoff. This is the case for $U \leq \widetilde{U}_E^k(\theta_E, p_E)$. Hence this condition is also sufficient for entry. The second part of the proposition shows how consumers of different types fare with entry. The rigid consumers never gain or lose because entry occurs only if the entrant can successfully attract the flexible consumers and, in that case, the incumbent has no reason to respond by offering the rigid consumers anything other than what it would offer in the absence of entry. However, the flexible consumers gain whenever there is entry because the entrant has to offer a higher utility to them to make it unattractive for the incumbent to more than match that offer.

5 Implications

5.1 Conditions for entry

We now use Proposition 2 to study three cases that highlight the role of organizational objectives in determining the conditions for entry. We focus on the case where the funding level is the same for both the incumbent and entrant, which applies, for example, to payment by diagnosis related group under US Medicare or payment by results in the English NHS.[7]

Proposition 3 *The following conditions for entry to occur apply.*

[7]We consider below whether differentiating the price between the incumbent and entrant is optimal.

1. *Both incumbent and potential entrant are for-profit providers.*

$$\hat{\gamma}^{FP}(\widetilde{U}_E^{FP}(\theta_E, p), \theta_I, p) = \min\left\{\frac{\theta_E}{\theta_I}, 1\right\}. \tag{16}$$

   A sufficient condition for entry is that $\theta_E \geq \theta_I$. For flexible consumers, entry increases observed quality but leaves unobserved quality at the minimal level.

2. *A not-for-profit incumbent competes with a for-profit potential entrant. A necessary condition for entry is that $\theta_E > \theta_I$. For flexible consumers, entry increases observed quality but reduces unobserved quality to the minimal level.*

3. *A not-for-profit incumbent competes with a not-for-profit potential entrant. A sufficient condition for entry is that $\theta_E > \theta_I$.*

This proposition shows how organizational form matters for entry conditions. In the two symmetric cases, both incumbent and entrant for-profit and both incumbent and entrant not-for-profit, an efficiency advantage ($\theta_E > \theta_I$) is sufficient for entry. However, in the asymmetric case, a not-for-profit incumbent with a for-profit entrant, an efficiency advantage is necessary but not sufficient for entry because the not-for-profit incumbent is able to provide unobserved quality.

To understand the implications of Proposition 3 in more detail, consider first the case in which both incumbent and potential entrant are for-profit providers (case 1). Proposition 3 establishes that an efficiency advantage for the entrant ($\theta_E \geq \theta_I$) is then sufficient (but not necessary) for entry. Even if $\theta_E < \theta_I$, entry is still possible if $\gamma \leq \theta_E/\theta_I$, since the incumbent may prefer to make a higher profit per consumer on just the rigid consumers than a lower profit per consumer on all consumers. With $\theta_E < \theta_I$, then $\widetilde{U}_E^{FP}(\theta_E, p) < \widetilde{U}_I^{FP}(\theta_I, p)$, so offering the consumer utility in the top line in (15) in Proposition 2 would impose a loss on the entrant. Hence, if entry occurs, it must be that the bottom line in (15) in Proposition 2 applies. Moreover, with a for-profit entrant, $u^{FP}(\theta_E, p) = 0$. Thus the utility of the flexible consumers is always $\overline{U}^{FP}(\gamma, \theta_I, p)$ defined in (13). This can be evaluated by equating $1-\gamma$ times the incumbent payoff in (2) for $U = 0$ to the incumbent payoff in (2) for $U = \overline{U}^{FP}(\gamma, \theta_I, p)$ to give

$$\overline{U}^{FP}(\gamma, \theta_I, p) = c^{-1}(\gamma p \theta_I).$$

But we know from Proposition 2 that the utility of flexible consumers is higher with entry and, because for-profit providers set unobservable quality at the minimal level, their observed quality must also be higher.

Now consider the case of a not-for-profit incumbent and a for-profit potential entrant (case 2 in Proposition 3) in which $\theta_E > \theta_I$ is necessary for entry.[8] For $\theta_E \leq \theta_I$,

[8] It is straightforward to derive expressions for $\widehat{\gamma}^{NP}(\widetilde{U}_E^k(\theta_E, p), \theta_I, p)$ for $k \in \{FP, NP\}$ for the cases in which the incumbent is a not-for-profit provider but these offer little additional insight and so are

the utility provided by a not-for-profit incumbent in the absence of entry, $U^*(\alpha, \theta_I, p)$, is greater than the highest utility a for-profit entrant with the same efficiency parameter can profitably provide, i.e. $\widetilde{U}_E^{FP}(\theta_E, p)$. That reflects the not-for-profit incumbent's provision of unobserved quality which gives it an implicit cost advantage when competing with a for-profit entrant. This cost advantage follows from the strict convexity of the cost function given that the for-profit entrant can provide utility only by expenditure on observable quality.[9] Thus it requires a more efficient potential entrant for entry to occur when the incumbent is not-for-profit than when the incumbent is for-profit. This may be a reason for the difficulty of obtaining effective for-profit competition in contexts such as the British NHS that have not-for-profit incumbents. There is, of course, an entrant efficiency level, θ_E, at which it is infeasible for the incumbent to compete with the entrant, i.e. for which the incumbent cannot feasibly offer $\widetilde{U}_E^{FP}(\theta_E, p)$ due to the breakeven constraint. At this point $\widehat{\gamma}^{NP}(\widetilde{U}_E^{FP}(\theta_E, p), \theta_I, p) = 1$ and there is entry for all $\gamma \in [0, 1]$. Thus, a large enough entrant efficiency advantage is sufficient for entry. Because the not-for-profit incumbent sets unobservable quality above the minimal level, whereas the for-profit entrant sets it at zero, unobservable quality for flexible consumers falls with entry. However, their utility increases with entry since observed quality increases.

The case in which both incumbent and potential entrant are not-for-profit providers (case 3 in Proposition 3) has similarities with the case of competing for-profit providers (case 1). Specifically, entry is possible with $\theta_E \leq \theta_I$ if γ is low enough and entry occurs for sure if $\theta_E > \theta_I$. This is because a not-for-profit entrant's choice of observable quality effectively commits it to providing the quality bundle that maximizes consumer utility subject to its breakeven constraint if that is required to attract flexible consumers. Thus, in contrast to entry by a for-profit provider, the incumbent no longer has an implicit cost advantage from its provision of unobserved quality. As a result, for $\theta_E > \theta_I$, then $\widehat{\gamma}^{NP}(\widetilde{U}_E^{NP}(\theta_E, p), \theta_I, p) = 1$ and entry occurs for all γ. In effect the incumbent cannot then offer the utility that a more efficient entrant offers to flexible consumers.

Comparing the conditions in cases 2 and 3 in Proposition 3, there is a range of $\theta_E > \theta_I$ for which there is no entry with a for-profit entrant while there is entry with a not-for-profit entrant. Thus entry can occur with a lower entrant efficiency advantage if the entrant is not-for-profit than if it is for-profit. Another way to think about this is that the range of γ for which there is entry when the entrant, as well as the incumbent, is not-for-profit is strictly wider than when the entrant is for-profit. This stems from the fact that the not-for-profit entrant enjoys a similar implicit cost advantage to the incumbent because its preferences ensure that it delivers positive unobservable qual-

not included in the proposition.

[9]This is different from Lakdawalla and Philipson (2006) where the cost advantage of a not-for-profit comes from its access to donations.

ity. It is consistent with much competition in education and health services in practice being by not-for-profit providers.

With competition between two not-for-profit providers, entry is not just about the cost advantage with the possibility of pure *paternalism induced entry*, that is, entry by a not-for-profit provider that occurs only because α is strictly greater than 1. To illustrate this, consider what happens if $\theta_E = \theta_I$, in which case neither incumbent nor entrant has an inherent cost advantage. We know from case 2 of Proposition 3 that a for-profit provider never enters in this case, so entry is possible only when there is a not-for-profit entrant. If α were equal to 1, the incumbent would, even without entry, always make the choices optimal for consumers given the breakeven constraint, so we would have $u^{NP}(\theta_I, p) = \tilde{U}_I^{NP}(\theta_I, p)$. An entrant with $\theta_E = \theta_I$ could not offer utility greater than this to attract flexible consumers, so entry would not occur. Thus, entry can occur in this case only because $\alpha > 1$. Moreover, $\tilde{U}_E^{NP}(\theta_E, p) = \tilde{U}_I^{NP}(\theta_I, p)$ when $\theta_E = \theta_I$. Thus, when flexible consumer utility is given by the upper line in (15) in Proposition 2, it is the highest the entrant can provide given the breakeven constraint. So the paternalism of the entrant is completely undone because flexible consumers get their maximal utility given the productive efficiency of the provider. This Bertrand-style competition leaves no scope for providers to prioritize their quality preferences over those of flexible consumers.

5.2 Comparative Statics

We now look at how the cost difference and the extent of the incumbent's paternalism affect the likelihood of entry. Here, we have the following result:

Proposition 4 *Suppose a not-for-profit incumbent competes with a for-profit potential entrant. Then the critical value of γ at which entry occurs, $\widehat{\gamma}^{NP}(\widetilde{U}_E^{FP}(\theta_E, p), \theta_I, p)$, is increasing in α and θ_E where it is less than one. If, moreover, there is entry when $\widetilde{U}_E^{FP}(\theta_E, p) < \widetilde{U}_I^{NP}(\theta_I, p)$, the utility of flexible consumers is increasing in α.*

A higher critical value of γ increases the range of γ for which entry takes place. Proposition 4 thus implies that a more paternalistic incumbent and a more efficient entrant increase the probability of entry. This makes intuitive sense. If the incumbent is more paternalistic then it would lean towards serving only the rigid consumers rather than compromising and serving the flexible consumers when the entrant tries to attract them. When θ_E is higher, the entrant can make a more aggressive offer to the flexible consumers in order to attract them.

Proposition 4 also considers the effect of α on the utility of flexible consumers conditional on entry. When $\widetilde{U}_E^{FP}(\theta_E, p) < \widetilde{U}_I^{NP}(\theta_I, p)$, the lower line in (15) in Proposition 2 applies, so flexible consumers receive utility $\overline{U}^{NP}(\gamma, \theta_I, p)$, and having a more paternalistic incumbent increases the value of entry to them. That is not the case when the upper line in (15) applies because then flexible consumers receive the highest utility the incumbent can afford, which is independent of α.

5.3 Consumer Utility and Organizational Objectives

Proposition 2 can also be used to compare the payoffs to consumers under competition when the incumbent is a not-for-profit provider, rather than a for-profit provider. Proposition 1 showed that, in the absence of competition, consumers are always better off with a not-for-profit provider no matter how much more efficient a for-profit provider is. The following result applies when there is competition.

Proposition 5 *With competition between providers, rigid consumers receive higher utility with a not-for-profit than with a for-profit incumbent. Flexible consumers receive higher utility with a not-for-profit incumbent than with a for-profit incumbent of the same efficiency; specifically,* $\tilde{U}_I^{NP}(\theta_I, p) \geq \bar{U}^{NP}(\gamma, \theta_I, p) > \tilde{U}_I^{FP}(\theta_I, p) \geq \bar{U}^{FP}(\gamma, \theta_I, p)$ *for all* (γ, θ_I, p).

This result establishes that, provided entry occurs, all consumers have higher utility with a not-for-profit incumbent than with a for-profit incumbent of the same efficiency. Of particular interest is the observation that a not-for-profit incumbent is always willing to offer higher utility to attract flexible consumers than the highest utility a for-profit incumbent of the same efficiency can afford. This is because the cost function is strictly convex, so a not-for-profit (which values unobservable quality) can provide given consumer utility at lower cost than a for-profit with the same efficiency parameter.

However, flexible consumers are not necessarily better off with a not-for-profit incumbent since, as shown by Proposition 3, a higher efficiency entrant is required for entry to occur than with a for-profit incumbent of the same efficiency. Thus entry may not occur with a not-for-profit incumbent even though it would have occurred with a for-profit incumbent of the same efficiency. In that case, flexible consumers may have lower utility with a not-for-profit incumbent than with a for-profit one. Rigid consumers, though, always do better with a not-for-profit incumbent, so the two types of consumers may have conflicting interests.

As regards different types of entrants, flexible consumers may also receive higher utility from having a not-for-profit entrant than from having a for-profit entrant with the same efficiency. This can happen in two ways. One is that a sufficiently productive not-for-profit entrant may choose to provide utility higher than the minimum required to attract flexible consumers, as a result of which the second term in the maximum expressions in (15) exceeds the first. In contrast, a for-profit entrant never offers utility higher than required to attract flexible consumers because to do so would lower profit. The other can occur even when that is not the case. Because the cost function is strictly convex and the not-for-profit entrant provides both types of quality, the highest utility that a not-for-profit entrant can afford is always strictly greater than the highest utility that a for-profit entrant with the same efficiency can afford, that is, $\tilde{U}_E^{NP}(\theta_E, p) > \tilde{U}_E^{FP}(\theta_E, p)$. Thus, from (14), a not-for-profit potential entrant may enter when a for-profit potential entrant with the same efficiency would not, which is consistent with much competition in education and health services being by not-for-profit providers. In this case, there is no conflict between rigid and flexible consumers.

6 Internalities and Externalities from Choice

In this section, we consider the implications of our framework for two of the main concerns that have been expressed about why competition may not raise the welfare of consumers. The first applies to flexible consumers and is based on an *internality*, i.e. the possibility that they do not exercise choice in their own best interest. The second applies to rigid consumers and is based on an *externality*, where the concern is that decisions made by flexible consumers are not in the best interest of rigid consumers. In each case, we offer a simple formalization where the incumbent is not-for-profit and the funding level is the same for all providers, i.e. $p_I = p_E = p$.

An Internality for Flexible Consumers The internality we consider is the classic one studied in the behavioral economics literature when consumers do not understand what is good for them as in, for example, Herrnstein et al (1993). An internality is like a within-person externality where consumers make choices that do not correspond to their true welfare. To model this, we suppose the preferences of the not-for-profit incumbent reflect consumers' true welfare. This corresponds to the common presumptions that teachers know better than parents how to organize a curriculum and that doctors should be guardians of treatment options in the provision of medical care.[10]

In that formulation consumers' true welfare is given by the incumbent's objective function $\alpha Q + q$ but consumers make their decisions according to what we here call their *decision preferences* $Q + q$. Whether entry occurs is then governed by the decision preferences of flexible consumers. It is not difficult to see that entry may then be excessive because flexible consumers switch to the entrant even when their true welfare is reduced. The next proposition gives sufficient conditions for this.

Proposition 6 *With an internality in consumer choice, a not-for-profit incumbent and $u^k(\theta_E, p) \leq \overline{U}^{NP}(\gamma, \theta_I, p)$ for entrant type $k \in \{FP, NP\}$, flexible consumer welfare is increased by prohibiting entry for any $\alpha > 1$.*

The essence of this result is that competition induces providers to compete by offering higher payoffs evaluated in terms of consumers' decision preferences. Thus, when

[10]The concern in some of the health economics literature about the impact of competition on quality can be thought of in terms of policy-makers having a paternalistic objective function.

the incumbent's preferences are a better representation of consumers' true welfare than consumers' decision preferences, flexible consumers may lose out from competition. By Proposition 2, the outcome for rigid consumers is unaffected by entry, so overall consumer welfare is also reduced. Note that the condition $u^k(\theta_E, p) \leq \overline{U}^{NP}(\gamma, \theta_I, p)$ is always satisfied when the entrant is for-profit because $u^{FP}(\theta_E, p) = 0$. But it also holds with a not-for-profit entrant whose costs are at least as high as the incumbent's ($\theta_E \leq \theta_I$).

This makes sense since the gains to flexible consumers from entry are illusory because they do not know their own welfare. Thus, the paternalism that we previously attributed to providers is now embraced by society (in the form of the policy objective) and hence there is no gain from permitting competition and the choice it brings. It is important to note that the result in Proposition 6 applies to entry by a not-for-profit entrant even if that entrant had the same preferences as the incumbent. This is because the entrant competes for flexible consumers in terms of their decision preferences, which limits the scope for it to exercise its paternalistic preferences. Thus, competition may, and for $u^{NP}(\theta_E, p) \leq \overline{U}^{NP}(\gamma, \theta_I, p)$ certainly will, undermine paternalism even when it is welfare improving and all providers share the same paternalistic preference.[11] But with a not-for-profit entrant that is sufficiently efficient to make $u^{NP}(\theta_E, p)$ high enough, entry may increase consumer welfare. Then, rather than prohibit entry, it is better for the policy-maker to restrict entry to not-for-profits and offer entrants a payment p_E sufficiently much lower than the payment p_I to the incumbent to deter entry by not-for-profits who are not sufficiently efficient to increase consumer welfare, see Section 7 below.

This section also has implications for the political economy of reform. If incumbent providers are speaking for consumers' true welfare, this puts a very different gloss on their opposition to entry discussed in Section 8 below. This is an important judgement for policy-makers to make when designing a system for providing public services.

Proposition 6 considers an extreme case in which the incumbent's preferences exactly match the true welfare of consumers. That is sufficient for flexible consumers to lose out from entry when $u^k(\theta_E, p) \leq \overline{U}^{NP}(\gamma, \theta_I, p)$ but it is certainly not necessary. The same conclusion can occur as long as the incumbent's preferences better represent the true welfare of consumers than consumers' own decision preferences.

An Externality for Rigid Consumers We now consider what happens when there is damage to rigid consumers when flexible consumers switch from consuming with the incumbent. A range of examples have been put forward in the literature on public service provision. One key possibility is that there is a positive "peer benefit" for rigid

[11]It is important to observe that, even though providers care about "true" consumer welfare, they are not pure altruists as they care about this only when they are actually serving the consumers. Thus, they are "warm glow" altruists.

consumers from having the flexible consumers served by the incumbent. This could be true when in education bright kids are more likely to have parents who shop around. It could also be important in medical care when the "voice" of flexible consumers is lost when they no longer use a service.[12]

We focus on the case where the incumbent is a not-for-profit, the entrant for-profit, and suppose that when all consumers are with the incumbent their payoff is $U^* (\alpha, \theta_I, p)$, whereas with exit of the flexible consumers the payoff of rigid consumers is $\nu U^* (\alpha, \theta_I, p)$, where $\nu < 1$ captures the cost to rigid consumers of having the flexible consumers switch to the entrant. If the proportion γ of flexible consumers is large enough, the gains to flexible consumers from entry will outweigh the losses to rigid consumers, so aggregate consumer welfare will increase. To illustrate the opposite possibility, consider γ sufficiently small that the utility of flexible consumers when switching to an entrant is given by the lower line in (15) in Proposition 2. Then the aggregate gain to consumers from entry is

$$\gamma \bar{U}^{NP}(\gamma, \theta_I, p) + [(1-\gamma)\nu - 1]U^*(\alpha, \theta_I, p). \tag{17}$$

This gain is clearly increasing in ν.

Proposition 7 *Suppose the incumbent is not-for-profit and the entrant for-profit. For small enough γ, there exists $\hat{\nu} < 1$ such that aggregate consumer welfare is decreasing with entry for all $\nu \leq \hat{\nu}$.*

Thus, if the externality is large enough, entry is no longer desirable. The condition in Proposition 7 could in principle be used to create a concrete welfare analysis based on the fraction of flexible consumers in the population, γ, and the amount of damage that their exit does to rigid consumers given the underlying gains that flexibility yields.

7 Pricing, Vouchers and Entry Policy

In this section, we explore the optimal funding level for the public service, including whether it is optimal to pay a per capita amount to an entrant different from that to the incumbent in order to encourage or discourage entry. With standard voucher schemes for education, such as that introduced in Sweden in 1992, a consumer can transfer the public funding to the entrant.[13] However, the value of a voucher could be different from the public funding for the incumbent and it could also be designed to offset the

[12]Peer effects are just one way in which consumers leaving a provider can have a negative externality on those who remain. Other obvious examples are when the incumbent has increasing returns to scale and when the entrant "cherry picks" consumers who cost less to serve. However, capturing these formally would require a somewhat different model.

[13]The kind of voucher that we have in mind here is like that used in Sweden where no consumer-financed "top-up" is allowed.

effects of externalities and internalities. Here we consider payment that is optimal from the perspective of consumers who pay taxes to fund the service with a constant marginal cost of public funds $\xi \geq 1$ and show that, in general, it is optimal to treat entrants and incumbents differently.[14] This in turn affects the probability of entry.

Monopoly Benchmark We begin with optimal per capita payment p_I for the public service for a not-for-profit incumbent in the absence of competition. A not-for-profit monopolist provides utility $U^*(\alpha, \theta_I, p_I)$ given by (7). Consumer welfare, taking taxation into account, is therefore

$$U^*(\alpha, \theta_I, p_I) - \xi p_I. \tag{18}$$

The first-order condition for the welfare-maximizing per capita payment p_I (assuming an interior solution) is

$$\frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial p_I} + \frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial p_I} = \xi.$$

The marginal benefit of an increase in resources available to the provider depends on the increase in each type of quality, while the cost is in the form of tax revenue raised to fund the service. Optimal funding is lower when the marginal cost of public funds is high. Since $\alpha > 1$, the quality mix provided by the not-for-profit incumbent is not optimal from the point of view of consumers, implying less generous funding than would be the case if $\alpha = 1$.

Entry In considering entry, we look at optimal funding from an ex ante perspective, i.e. before the efficiency of the potential entrant is known, and focus on the case of a not-for-profit incumbent facing a for-profit potential entrant. For given (p_I, p_E), there will then, by Proposition 2, be entry if θ_E is large enough. Specifically, let $\hat{\theta}_E(p_E, p_I)$ denote the entrant efficiency level that makes the incumbent just unwilling to offer consumer utility $\widetilde{U}_E^{FP}(\theta_E, p_E)$ defined by (3), the highest consumer utility a for-profit entrant is prepared to offer, to retain the flexible consumers. ($\hat{\theta}_E(p_E, p_I)$ also depends on θ_I but that is taken as fixed for this analysis.) This efficiency level is defined for-

[14] We could, as in standard models of regulation, introduce a welfare weight that values providers' payoffs, though possibly somewhat less than consumer utility. Our framework is, however, somewhat non-standard because, in the case of not-for-profit provision, provider payoffs take the form of "decision rents" rather than monetary profits. Moreover, the question of how the welfare of teachers and doctors should count in the provision of the services is moot. In political economy models, it is common to ignore the welfare of providers (politicians and bureaucrats) and simply count the welfare of voters. In the case of for-profit providers, the policy debate often proceeds as if there should be a negative weight on profit in public service provision. For example, in the UK, there is a campaign called "Public Services Not Private Profit" supported by around 14 major trade unions whose objective could be interpreted in this way, as could the objective of a lobby group such as "We Own it" https://weownit.org.uk/ whose strap line is "Public Services for People not Profit".

mally by

$$\gamma = \widehat{\gamma}^j(c^{-1}(p_E\hat{\theta}_E(p_E, p_I)), \theta_I, p_I). \tag{19}$$

The probability of entry is then $1 - G\left(\hat{\theta}_E(p_E, p_I)\right)$. Also let $\hat{U}(\theta_I, \gamma, p_I)$ be the utility of a flexible consumer who switches to the entrant as given by Proposition 2. It does not depend on either θ_E or p_E because $u^k(\theta_E, p_E) = 0$ for all (θ_E, p_E) for $k = FP$, so the consumer utilities in (15) do not in this case depend on (θ_E, p_E). Ex ante expected consumer welfare is then

$$\left[(1-\gamma) + \gamma G\left(\hat{\theta}_E(p_E, p_I)\right)\right]\left[U^*(\alpha, \theta_I, p_I) - \xi p_I\right] \\ + \gamma\left[1 - G\left(\hat{\theta}_E(p_E, p_I)\right)\right]\left[\hat{U}(\theta_I, \gamma, p_I) - \xi p_E\right]. \tag{20}$$

The first term is the welfare of rigid consumers plus that of flexible consumers for the entrant efficiency levels for which there is no entry, i.e. $\theta_E < \hat{\theta}_E(p_E, p_I)$, the second term the welfare of flexible consumers when $\theta_E \geq \hat{\theta}_E(p_E, p_I)$ and hence entry occurs. Changing the payments to providers has three main effects on welfare in (20). Increasing funding to either the incumbent or the entrant necessitates higher taxes which reduce consumers' welfare. Counteracting this is an increase in quality. For rigid or flexible consumers who remain with the not-for-profit incumbent, this effect is direct. However, increasing p_I also affects the utility of flexible consumers who switch since their utility level is set by what the incumbent would be prepared to offer to retain them. Third, funding arrangements change the probability of entry, i.e. the critical efficiency level at which an entrant finds it worthwhile to enter.

It follows from (20) that the *possibility* of entry affects the generosity with which the incumbent should be funded, i.e. it is no longer optimal to set p_I to maximize the expression in (18). This is because monopoly pricing ignores the effect of p_I on the probability and consequences of entry — a more generous funding level for the incumbent will tend to discourage entry and will also make flexible consumers better off when entry occurs. Thus, in general, the optimal per capita payment to the incumbent, p_I, is different when entry is permitted from when it is not. That said, more information is required to specify whether p_I should be higher or lower when entry is a possibility than when it is not.

We can, however, be more specific about the optimal payment to the entrant. To state the results on this, let $\Delta U(p_I) = \hat{U}(\theta_I, \gamma, p_I) - U^*(\alpha, \theta_I, p_I)$ be the utility gain to a flexible consumer of switching to the entrant.[15]

Proposition 8 *Suppose that a not-for-profit incumbent competes with a for-profit entrant and a policy-maker sets the per capita payment to the entrant, p_E, to maximize expected consumer welfare (20) for given per capita payment to the incumbent, p_I. Then for $g(\theta_E)$ log-concave*

[15] Arguments other than p_I are suppressed for notational simplicity because they are constant in the analysis of this section.

and optimal $\hat{\theta}_E(p_E, p_I) \in (\underline{\theta}, \bar{\theta})$*, the optimal per capita payment to the entrant is the unique* p_E^* *that satisfies*

$$\left[\frac{\Delta U(p_I)}{\xi} + (p_I - p_E^*)\right] \frac{\hat{\theta}_E(p_E^*, p_I)}{p_E^*} = \frac{1 - G(\hat{\theta}_E(p_E^*, p_I))}{g(\hat{\theta}_E(p_E^*, p_I))}. \tag{21}$$

Equation (21) applies for any p_I, including the optimal value that maximizes expected consumer welfare (20). In general, it implies $p_E^* \neq p_I$.[16] To understand the implications of Proposition 8, define

$$\eta(\hat{\theta}_E(p_E, p_I)) = \frac{g(\hat{\theta}_E(p_E, p_I))\,\hat{\theta}_E(p_E, p_I)}{1 - G(\hat{\theta}_E(p_E, p_I))} \tag{22}$$

as the elasticity of the entry probability with respect to the payment to the entrant. This depends on the shape of the distribution of the potential entrant's efficiency parameter, θ_E. Rearranging (21), we have the following formula for the optimal payment to the entrant

$$p_E^* = \frac{\eta(\hat{\theta}_E(p_E^*, p_I))\,[\Delta U(p_I) + \xi p_I]}{1 + \eta(\hat{\theta}_E(p_E^*, p_I))\,\xi}, \tag{23}$$

which also holds for *any* value of p_I. The value of the payment is thus increasing in $\eta(\cdot)$, i.e. the more responsive is entry to a higher payment then the larger it is all else equal. The payment should also be more generous when the marginal gain to the flexible consumers from switching to the entrant, $\Delta U(p_I)$, is larger. This makes sense as entry is better for flexible consumers in this case.

An attractive feature of (23) is that it depends on magnitudes that can be estimated in specific applications. For example suppose that the entrant efficiency parameter θ_E follows a Pareto distribution with shape parameter ζ, i.e. $G(\theta_E) = 1 - (\underline{\theta}/\theta_E)^{\zeta}$, then $\eta(\hat{\theta}_E(p_E, p_I)) = \zeta$. This is motivated by noting from Axtell (2001) that the size distribution of firms suggests that productivity follows a Zipf distribution, i.e. a Pareto distribution with $\zeta = 1$. A value of $\xi = 1.5$ is a reasonable figure in line with many estimates of the cost of public funds and p_I would be known from the funding levels currently used in the market. The only additional element of (23) that would be needed to apply the formula for policy purposes would be $\Delta U(p_I)$, i.e. the "willingness to pay" by flexible consumers to switch to the entrant.

To illustrate how to apply this formula, consider the case of hip replacement surgery in the UK. A National Health Service (NHS) provider is paid around £5000 per operation while the cost of private treatment is around £10,000. If the latter is all out of

[16] Log-concavity of the density is satisfied by many standard probability distributions and is widely used in economic models, see Bagnoli and Bergstrom (2005). Its role in Proposition 8 is to ensure that (21) has only one solution for p_E^*. Without it, the optimal entrant payment will still satisfy (21) but there may be other solutions as well, so one would have to check which corresponds to a true maximum of (20).

pocket, we could use it as a rough estimate $\Delta U(p_I)$ because it measures consumers' willingness to pay for the additional benefit of the private treatment. Then if $\eta = 1$ and $\xi = 1.5$, the optimal amount for the NHS to pay a private provider to offer its services for free to NHS patients should be

$$p_E^* = \frac{£10,000 + £5,500}{2.5} = £6200.$$

So this is a case where the per capita payment to the entrant should be larger than the current per capita payment to the incumbent. These specific numbers are, of course, only illustrative but they show how Proposition 8 can be applied to practical cases.[17]

Another nice feature of (23) is that it easily generalizes to situations where the gain for flexible consumers and/or loss to rigid consumers is affected by an internality or externality. Indeed, we simply replace $\Delta U(p_I)$ with a modified expression which takes this into account. Thus in the case of the internality in Section 6

$$\Delta U(p_I) = \hat{U}(\theta_I, \gamma, p_I) - [\alpha Q^*(\alpha, \theta_I, p_I) + q^*(\alpha, \theta_I, p_I)],$$

while for the externality in Section 6

$$\Delta U(p_I) = \hat{U}(\theta_I, \gamma, p) - \frac{[1-(1-\gamma)\nu]}{\gamma} U^*(\alpha, \theta_I, p_I).$$

It is easy to see that, when there are forces which lower the gain to the flexible consumers, entrants should receive less favorable treatment. In other words, pricing and voucher programs need to be "smart", taking into account any imperfections in the choice and competition mechanism. Banning entry is a blunt instrument compared to setting p_E optimally because optimal pricing allows entry to occur when an entrant has sufficiently low costs. But it may still be that restricting entry to not-for-profit providers is beneficial, as discussed in Section 6.

8 Political Economy

The normative aspects of the value of choice and competition we have focused on so far concern consumer interests. We now consider the possibility that provider interests weigh heavily in the policy-maker's decisions. This is best motivated as a political economy model where providers have influence. A natural way to micro-found this would be to use a menu auction model along the lines of Grossman and Helpman (2001) where providers can make transfers to policy makers to influence policy. It seems natural to suppose that incumbents would have better access to policy makers

[17] The argument presented here can be extended to cover the case where the entrant is not-for-profit firm. This will affect the critical $\hat{\theta}(p_E, p_I)$ but the core factors which shape optimal funding for entrants remain the same.

compared to potential entrants. In such models, policy-makers end up maximizing a weighted sum of provider and consumer welfare. We consider the implications of this for whether entry will be permitted.

We revert to the case with $p_E = p_I = p$. With δ_I and δ_E the weights on the incumbent's and the entrant's welfare respectively, ex ante expected welfare with provider welfare added to the consumer welfare in (20) becomes

$$\left[(1-\gamma)+\gamma G\left(\hat{\theta}_E\right)\right]\left[\delta_I v_I^j\left(u^j\left(\theta_I,p\right),\theta_I,p\right)+u^j\left(\theta_I,p\right)\right]-\xi p$$
$$+\gamma\int_{\hat{\theta}_E}^{\bar{\theta}}\left[\delta_E v_I^j(\hat{U}\left(\theta_I,\gamma,p\right),\theta_I,p)+\hat{U}\left(\theta_I,\gamma,p\right)\right]dG\left(\theta\right). \quad (24)$$

In our framework, it does not matter whether, in their craving for political influence, providers are motivated by paternalistic preferences or financial considerations. Moreover, the reason for provider paternalism is also not important. It could reflect a genuine concern for particular dimensions of quality or simply a desire to organize the service for the convenience of providers rather than in the interests of consumers. Whatever their motive, our framework suggests that all incumbent providers will tend to oppose entry, whether for-profit or not-for-profit. Moreover, the interests of insiders in provision are often powerful; teachers' and doctors' unions frequently have a powerful voice in policy debates which they use to oppose entry. Moreover, we would expect such opposition even if consumer welfare is increased by allowing entry and incumbents have not-for-profit motives.

We summarize this discussion as:

Proposition 9 *A policy-maker will oppose permitting choice and competition if δ_I is sufficiently large.*

This result echoes the literature on regulatory capture, an idea emphasized in particular by Stigler (1971) and analyzed more formally in Laffont and Tirole (1993) but applies equally well when insiders have not-for-profit objectives.

9 Concluding Comments

This paper has developed a framework for analyzing choice and competition in public service provision that recognizes the role for not-for-profit provision. Our approach to modeling not-for-profit provision combines the insight of Hansmann (1980) that such providers can be valuable when there is an unobservable dimension to their output with the recognition by Newhouse (1970) that many such providers have a bias towards quality that reflects producer interests but may also be paternalistic. The view that provider interests matter fits a range of services where physicians, lawyers and teachers run public services according to their views of what is good for consumers which means that providers earn decision rents even if they have not-for-profit status. Monopoly provision with public funding is then no guarantee that consumers get what they want from public services even if incumbents provide some unobserved quality which would not be provided by profit-maximizing firms. There is then a role for competition agenda beyond concerns about cost efficiency. The model generates the following substantive findings.

First, with monopoly provision, consumers are better off with a not-for-profit provider no matter how much more efficient an alternative for-profit provider might be. This is consistent with many local services being provided on a not-for-profit basis. Entry by either a for-profit or a not-for-profit provider benefits consumers who switch to the entrant but with observable dimensions of quality improved at the expense of unobservable ones — in the latter case, competition reduces the scope for the exercise of paternalism by the entrant. However, entry by a not-for-profit provider requires less cost advantage over the incumbent than entry by a for-profit provider, which is consistent with much competition in education and health services being by not-for-profit providers.

Second, when entry occurs, all consumers are better off with a not-for-profit incumbent than with an equally efficient for-profit incumbent. Then, it is beneficial for consumers that a not-for-profit provider remains in business. However, entry requires less efficiency advantage if the incumbent is for-profit than if it is not-for-profit. So there is a trade-off: with a for-profit incumbent, consumers are more likely to get the benefit of entry but, with a not-for-profit incumbent, consumers are better off provided that entry occurs.

Third, paternalism may not always be misplaced. Consumers may, for example, not fully appreciate the value of some dimensions of quality. It is then not surprising that entry may reduce welfare compared to a paternalistic not-for-profit incumbent. In such circumstances, there may be a case for restricting entry to not-for-profit providers.

Fourth, we show that with a voucher scheme which permits entry, a "level playing field" is not generally optimal. Depending on ex ante market conditions, it may be optimal to pay the entrant either less or more than the incumbent. The model offered an insight into the factors which determine optimal vouchers that could be applied in specific situations.

Fifth, even when the provider is a not-for-profit firm, producer interests will tend to oppose entry if these reduce decision rents of providers. Thus there are likely to be provider-consumer tensions in markets that are opened up to competition even when the provider does not care about profit. Monopoly rents in our framework are consumed in the form of provider paternalism. Policy-makers who are contemplating opening up services to competition should be cognizant of this bias when deciding whether to do so.

References

[1] Axtell, Robert L. [2001], "Zipf distribution of U.S. firm sizes," *Science*, Vol. 293, Issue 5536, 1818-1820.

[2] Bagnoli, Mark and Ted Bergstrom [2005], "Log-concave probability and its applications," *Economic Theory*, 26, 445-469.

[3] Barseghyan, Levon, Damon Clark and Stephen Coate [2014], "Public school choice: an economic analysis," unpublished typescript.

[4] Besley, Timothy and Maitreesh Ghatak [2001], "Government versus private ownership of public goods," *Quarterly Journal of Economics*, 116(4), 1343-1372.

[5] Besley, Timothy and Maitreesh Ghatak [2005], "Competition and incentives with motivated agents," *American Economic Review*, 95(3), 616-636.

[6] Böhlmark, Anders and Mikael Lindahl, [2015], "Independent schools and long-run educational outcomes: evidence from Sweden's large-scale voucher reform," *Economica*, 82, 508-551.

[7] Brekke, Kurt R., Luigi Siciliani and Odd Rune Straume, [2011], "Hospital competition and quality with regulated prices," *Scandinavian Journal of Economics*, 113(2), 444-469.

[8] Brekke, Kurt R., Luigi Siciliani and Odd Rune Straume, [2012], "Quality competition with profit constraints," *Journal of Economic Behavior and Organization*, 84, 642-659.

[9] Brekke, Kurt R., Rosella Levaggi, Luigi Siciliani and Odd Rune Straume, [2014], "Patient mobility, health care quality and welfare," *Journal of Economic Behavior and Organization*, 105, 140-157.

[10] Card, David, Martin D. Dooley and A. Abigail Payne, [2010], "School competition and efficiency with publicly funded Catholic schools," *American Economic Journal: Applied Economics*, 2(4), 150-176.

[11] Chalkley, Martin and James M. Malcomson [2000], "Government purchasing of health services," in Anthony J. Culyer, and Joseph P. Newhouse (eds), *Handbook of Health Economics*, Amsterdam: Elsevier.

[12] Chandra, Amitabh, Amy Finkelstein, Adam Sacarny and Chad Syverson [2016], "Health care exceptionalism? Performance and allocation in the US health care sector," *American Economic Review*, 106(8), 2110-2144.

[13] Epple, Dennis and Richard Romano [1998], "Competition between private and public schools, vouchers, and peer-group effects," *American Economic Review*, 88(1), 33-62.

[14] Friedman, Milton [1962], *Capitalism and Freedom*, Chicago: University of Chicago Press.

[15] Gaynor, Martin, Kate Ho and Robert J. Town [2015], "The industrial organization of health-care markets," *Journal of Economic Literature*, 53(2), 235-284.

[16] Gaynor, Martin, Carol Propper and Stephen Seiler [2012], "Free to choose: reform and demand response in the English National Health Service," NBER Working Paper 18574.

[17] Ghatak, Maitreesh and Hannes Mueller, [2011], "Thanks for nothing? Not-for-profits and motivated agents," *Journal of Public Economics,* 95 (1), 94-105.

[18] Gibbons, Stephen, Stephen Machin and O. Silva [2008], "Choice, competition, and pupil achievement," *Journal of the European Economic Association*, 6(4), 912-947.

[19] Glaeser, Edward and Andrei Shleifer [2001], "Not-for-profit entrepreneurs," *Journal of Public Economics*, 81, 99-115.

[20] Grossman, Gene and Elhanan Helpman [2001], *Special Interest Politics*, Cambridge MA: MIT Press.

[21] Hansmann, Henry [1980], "The role of nonprofit enterprise," *Yale Law Journal*, 89, 835-901.

[22] Hart, Oliver, Andrei Shleifer and Robert Vishny [1997], "The proper scope of government: theory and an application to prisons," *Quarterly Journal of Economics*, 108, 1119-1158.

[23] Herrnstein, R.J., George F. Loewenstein, Drazen Prelec and William Vaughan [1993], "Utility maximization and melioration: Internalities in individual choice,"*Journal of Behavioral Decision Making*, 6, 149-185.

[24] Hoxby, Caroline M. [1999], "The productivity of schools and other local public goods," *Journal of Public Economics*, 74(1), 1-30.

[25] Hoxby, Caroline M. [2003], "School choice and school productivity: should school choice be a tide that lifts all boats," in Caroline M. Hoxby (ed), *The Economics of School Choice*, NBER, University of Chicago Press.

[26] Jewitt, Ian [1987], "Risk aversion and the choice between risky prospects: the preservation of comparative statics results," *Review of Economic Studies*, 54, 73-86.

[27] Ladd, Helen F. [2002], "School vouchers: a critical view," *Journal of Economic Perspectives*, 16(4), 3-24.

[28] Laffont, Jean-Jacques and Jean Tirole [1993], *A Theory of Incentives in Procurement and Regulation*, Cambridge, MA: MIT Press.

[29] Laine, Liisa and Ching-to Albert Ma [2016], "Quality and competition between public and private firms," unpublished typescript.

[30] Lakdawalla, Darius and Tomas Philipson, [2006], "The nonprofit sector and industry performance," *Journal of Public Economics*, 90(9), 1681-1698.

[31] Lavy, Victor [2010], "Effects of free choice among public schools,"*Review of Economic Studies*, 77(3), 1164-1191.

[32] Le Grand, Julian [2007], *The Other Invisible Hand: Delivering Public Services through Choice and Competition*, Princeton NJ and Oxford: Princeton University Press.

[33] McMillan, Robert [2005], "Competition, incentives, and public school productivity," *Journal of Public Economics*, 89, 1131-1154.

[34] Neal, Derek [2002], "How vouchers could change the market for education," *Journal of Economic Perspectives*, 16(4), 25-44.

[35] Nechyba, Thomas J. [2000], "Mobility, targeting, and private-school vouchers," *American Economic Review*, 90(1), 130-146.

[36] Newhouse, Joseph P. [1970], "Toward a theory of nonprofit institutions: an economic model of a hospital," *American Economic Review*, 60(1), 64-74.

[37] Stigler, George [1971], "The theory of economic regulation," *Bell Journal of Economics and Management Science*, 2, 3-21.

[38] Tooley, James and Pauline Dixon, [2005], "Private education is good for the poor: a study of private schools serving the poor in low-income countries," Cato Institute, available at http://object.cato.org/sites/cato.org/files/pubs/pdf/tooley.pdf.

Appendix A Proofs

Proof of Proposition 1 That $u^{FP}(\theta_I, p_I) = 0$ for all (θ_I, p_I) follows directly from maximization of $v_I^{FP}(U, \theta_I, p_I)$ specified in (2) subject to $U \geq 0$. That $u^{NP}(\theta_I, p_I) = U^*(\alpha, \theta_I, p_I) > 0$ for all (θ_I, p_I) follows from the definition of $U^*(\alpha, \theta_I, p_I)$ in (7) and that this is strictly positive because both $Q^*(\alpha, \theta_i, p_i)$ and $q^*(\alpha, \theta_i, p_i)$ are strictly positive. With a monopoly not-for-profit incumbent, consumers' utility is given by (7) with $i = I$. For any parameter $z \in \{\alpha, \theta_I, p_I\}$,

$$\frac{\partial U^*(\alpha, \theta_I, p_I)}{\partial z} = \frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial z} + \frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial z}. \tag{A.1}$$

From the first-order conditions (6), note that μ must be strictly greater than zero, so the profit constraint (5) holds with equality. From these, for $i = I$ and $x_I > 0$,

$$\alpha c'(q^*(\alpha, \theta_I, p_I)) = c'(Q^*(\alpha, \theta_I, p_I)) \tag{A.2}$$

and

$$\theta_I p_I = c(Q^*(\alpha, \theta_I, p_I)) + c(q^*(\alpha, \theta_I, p_I)). \tag{A.3}$$

Consider first $z = \alpha$. Differentiation of (A.3) with respect to α gives

$$c'(Q^*(\alpha, \theta_I, p_I)) \frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial \alpha} + c'(q^*(\alpha, \theta_I, p_I)) \frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial \alpha} = 0$$

and, hence,

$$\frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial \alpha} = -\frac{c'(Q^*(\alpha, \theta_I, p_I))}{c'(q^*(\alpha, \theta_I, p_I))} \frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial \alpha}. \tag{A.4}$$

Substitution for $c'(Q^*(\alpha, \theta_I, p_I))$ in (A.4) from (A.2) gives

$$\frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial \alpha} = -\alpha \frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial \alpha}$$

which, substituted into (A.1) for $z = \alpha$, gives

$$\frac{\partial U^*(\alpha, \theta_I, p_I)}{\partial \alpha} = (1 - \alpha) \frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial \alpha}.$$

Differentiation of (A.2) with respect to α gives

$$\alpha c''(q^*(\alpha, \theta_I, p_I)) \frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial \alpha} + c'(q^*(\alpha, \theta_I, p_I)) \\ - c''(Q^*(\alpha, \theta_I, p_I)) \frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial \alpha} = 0.$$

Substitution for $\partial q^*(\alpha, \theta_I, p_I)/\partial\alpha$ in this from (A.4) gives

$$\frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial \alpha}\left[\frac{c'(Q^*(\alpha, \theta_I, p_I))}{c'(q^*(\alpha, \theta_I, p_I))}\alpha c''(q^*(\alpha, \theta_I, p_I)) + c''(Q^*(\alpha, \theta_I, p_I))\right] = c'(q^*(\alpha, \theta_I, p_I)),$$

which implies $\partial Q^*(\alpha, \theta_I, p_I)/\partial\alpha > 0$ and hence $\partial U^*(\alpha, \theta_I, p_I)/\partial\alpha < 0$ because c is strictly increasing and strictly convex and $\alpha > 1$.

Consider now $z = \theta_I$. Differentiation of (A.2) and (A.3) with respect to θ_I gives

$$\alpha c''(q^*(\alpha, \theta_I, p_I))\frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial \theta_I} - c''(Q^*(\alpha, \theta_I, p_I))\frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial \theta_I} = 0$$

and

$$c'(Q^*(\alpha, \theta_I, p_I))\frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial \theta_I} + c'(q^*(\alpha, \theta_I, p_I))\frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial \theta_I} = p_I. \tag{A.5}$$

The former can be solved for $\partial Q^*(\alpha, \theta_I, p_I)/\partial\theta_I$ to give

$$\frac{\partial Q^*(\alpha, \theta_I, p_I)}{\partial \theta_I} = \frac{\alpha c''(q^*(\alpha, \theta_I, p_I))}{c''(Q^*(\alpha, \theta_I, p_I))}\frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial \theta_I}. \tag{A.6}$$

Use of this in (A.1) for $z = \theta_I$ gives

$$\frac{\partial U^*(\alpha, \theta_I, p_I)}{\partial \theta_I} = \left[\frac{\alpha c''(q^*(\alpha, \theta_I, p_I))}{c''(Q^*(\alpha, \theta_I, p_I))} + 1\right]\frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial \theta_I},$$

which is positive if $\partial q^*(\alpha, \theta_I, p_I)/\partial\theta_I > 0$. Use of (A.6) in (A.5) and substitution for $c'(Q^*(\alpha, \theta_I, p_I))$ from (A.2) gives

$$\left[\frac{\alpha^2 c''(q^*(\alpha, \theta_I, p_I))}{c''(Q^*(\alpha, \theta_I, p_I))} + 1\right]c'(q^*(\alpha, \theta_I, p_I))\frac{\partial q^*(\alpha, \theta_I, p_I)}{\partial \theta_I} = p_I,$$

from which $\partial q^*(\alpha, \theta_I, p_I)/\partial\theta_I > 0$ because c is strictly increasing and strictly convex. For $z = p_I$, the argument is essentially identical to that for $z = \theta_I$.

Proof of Proposition 2 Suppose entry were to occur when (14) does not hold. Then, by the definition of $\widehat{\gamma}^j(U, \theta_I, p_I)$ in (12), the incumbent would compete to supply the whole market for even the highest payoff $\widetilde{U}_E^k(\theta_E, p_E)$ the entrant would be willing to offer the γ flexible consumers. So the entrant would not succeed in acquiring the flexible consumers and thus no entry would occur, which is a contradiction.

Suppose now (14) holds. Then, by the definition of $\widehat{\gamma}^j(U, \theta_I, p_I)$ in (12), the incumbent would not compete for the γ flexible consumers if the entrant were to offer $\widetilde{U}_E^k(\theta_E, p_E)$. By offering $\widetilde{U}_E^k(\theta_E, p_E)$, the entrant would make no less payoff than from not entering and would acquire the flexible consumers, so entry occurs.

For determining consumer payoffs, there are two cases to consider.

Case 1: $\gamma \geq \widehat{\gamma}^j(\widetilde{U}_I^j(\theta_I, p_I), \theta_I, p_I)$. In this case, there are sufficient flexible consumers for it to be worth the incumbent competing for them at the highest utility it is ever prepared to offer, $\widetilde{U}_I^j(\theta_I, p_I)$. If entry occurs (which in this case is only if $\widetilde{U}_E^k(\theta_E, p_E) \geq \widetilde{U}_I^j(\theta_I, p_I)$ because otherwise (14) is not satisfied), the entrant offers utility of $\widetilde{U}_I^j(\theta_I, p_I)$ so that it is not worth the incumbent attracting flexible consumers or, if higher, the payoff $u^k(\theta_E, p_E)$ it would offer in the absence of competition. The incumbent offers utility $u^j(\theta_I, p_I)$ and attracts only the rigid consumers, who thus receive utility $u^j(\theta_I, p_I)$. Flexible consumers choose the entrant and receive payoff $\max\{\widetilde{U}_I^j(\theta_I, p_I), u^k(\theta_E, p_E)\}$.

Case 2: $\gamma < \widehat{\gamma}^j(\widetilde{U}_I^j(\theta_I, p_I), \theta_I, p_I)$. In this case, there are insufficient flexible consumers for it to be worth the incumbent competing for them at the highest utility it is ever prepared to offer, $\widetilde{U}_I^j(\theta_I, p_I)$. If entry occurs, therefore, the entrant offers the lowest consumer payoff, $\overline{U}^j(\gamma, \theta_I, p_I)$ defined in (13), for which it is not worth the incumbent competing for flexible consumers or, if higher, the payoff $u^k(\theta_E, p_E)$ it would offer in the absence of competition. The incumbent then offers $u^j(\theta_I, p_I)$ and serves only the rigid consumers, who thus receive utility $u^j(\theta_I, p_I)$. Flexible consumers choose the entrant and receive payoff $\max\{\overline{U}^j(\gamma, \theta_I, p_I), u^k(\theta_E, p_E)\}$.

Entry increases the utility of flexible consumers because $\tilde{U}_I^j(\theta_I, p_I) \geq \overline{U}^j(\gamma, \theta_I, p_I) > u^j(\theta_I, p_I)$ and leaves utility of rigid consumers unchanged because they receive $u^j(\theta_I, p_I)$ both with and without entry.

Proof of Proposition 3 **Case 1:** For a for-profit incumbent with $p_I = p$, $u^{FP}(\theta_I, p) = 0$. Use of this and the incumbent payoff, (2) for $i = I$, in (12) gives

$$\hat{\gamma}^{FP}(U, \theta_I, p) = \min\left\{\frac{c(U) - c(0)}{p\theta_I - c(0)}, 1\right\}, \tag{A.7}$$

which yields $\hat{\gamma}^{FP}(U, \theta_I, p) = 1$ only if $U \geq \tilde{U}_I^{FP}(\theta_I, p)$. When the potential entrant is also a for-profit provider, from (3) for $i = E$, $\tilde{U}_E^{FP}(\gamma, \theta_E)$ satisfies

$$c(\tilde{U}_E^{FP}(\theta_E, p)) = p\theta_E. \tag{A.8}$$

Use of this and $c(0) = 0$ in (A.7) gives (16). By Proposition 2, entry occurs if $\gamma \leq \widehat{\gamma}^j(\widetilde{U}_E^k(\theta_E, p), \theta_I, p)$. From (16), when $\theta_E \geq \theta_I$, then $\widehat{\gamma}^j(\widetilde{U}_E^k(\theta_E, p), \theta_I, p) = 1$, so entry occurs for any $\gamma \leq 1$. For-profit providers always set $Q_I = Q_E = 0$, so unobserved quality for flexible consumers is the same minimal level with entry as without. But, from Proposition 2, their utility increases with entry so it must be that $q_E > q_I$. Thus entry increases observable quality for flexible consumers.

Case 2: For a not-for-profit incumbent with $p_I = p$, $u^{NP}(\theta_I, p) = U^*(\alpha, \theta_I, p)$. We first show that, if $\theta_E \leq \theta_I$, then $U^*(\alpha, \theta_I, p) > \tilde{U}_E^{FP}(\theta_E, p)$. Suppose not. Then from (3),

$$\begin{aligned} p\theta_E &= c\left(\tilde{U}_E^{FP}(\theta_E, p)\right) \\ &> c\left(\tilde{U}_E^{FP}(\theta_E, p) - q^*(\alpha, \theta_I, p)\right) + c\left(q^*(\alpha, \theta_I, p)\right) \\ &> c\left(U^*(\alpha, \theta_I, p) - q^*(\alpha, \theta_I, p)\right) + c\left(q^*(\alpha, \theta_I, p)\right) \\ &= p\theta_I. \end{aligned}$$

The second line of this follows because c is strictly convex, the third line from the supposition that $U^*(\alpha, \theta_I, p) \leq \tilde{U}_E^{FP}(\theta_E, p)$, and the final line because the breakeven constraint for the not-for-profit incumbent (5) holds with equality. This contradicts $\theta_E \leq \theta_I$ so it must be the case that, when that condition holds, $U^*(\alpha, \theta_I, p) > \tilde{U}_E^{FP}(\theta_E, p)$. But then the payoff to consumers that the not-for-profit incumbent would choose to offer even if not competing for flexible consumers is greater than the highest payoff the for-profit potential entrant would offer them. So the entrant would never attract the flexible consumers and so would not enter. Since in this case $Q_I > 0$ and $Q_E = 0$, entry reduces unobservable quality for flexible consumers to the minimal level. But, from Proposition 2, their utility increases with entry so it must be that $q_E > q_I$. Thus entry increases observable quality for flexible consumers.

Case 3: For $\theta_E > \theta_I$, $\tilde{U}_E^{NP}(\theta_E, p) > \tilde{U}_I^{NP}(\theta_I, p)$. The entrant is, therefore, always willing to offer utility higher than the incumbent can afford to attract flexible consumers, so entry always occurs.

Proof of Proposition 4 Let

$$\varphi(U, \alpha, \theta_I, p) = \frac{\alpha\hat{Q}^*(U, \theta_I, p) + \hat{q}^*(U, \theta_I, p)}{\alpha Q^*(\alpha, \theta_I, p) + q^*(\alpha, \theta_I, p)}.$$

Observe that the denominator in this is a maximum value function with $Q^*(\alpha, \theta_I, p)$ and $q^*(\alpha, \theta_I, p)$ the maximizers and that α enters only the objective function and not the constraints. So, by the envelope theorem, its derivative with respect to α is just $Q^*(\alpha, \theta_I, p)$. Moreover, $\hat{Q}^*(U, \theta_I, p)$ and $\hat{q}^*(U, \theta_I, p)$ are independent of α. Thus

$$\begin{aligned} \operatorname{sgn} \frac{\partial\varphi(U, \alpha, \theta_I, p)}{\partial\alpha} = \operatorname{sgn} \quad &\left(\hat{Q}^*(U, \theta_I, p)\, q^*(\alpha, \theta_I, p)\right. \\ &\left. - Q^*(\alpha, \theta_I, p)\, \hat{q}^*(U, \theta_I, p)\right) < 0, \end{aligned} \tag{A.9}$$

the inequality following because $\hat{Q}^*(U, \theta_I, p) < Q^*(\alpha, \theta_I, p)$ and $\hat{q}^*(U, \theta_I, p) > q^*(\alpha, \theta_I, p)$. Note from (12) that, for entry to occur with $\widehat{\gamma}^j(U, \theta_I, p) = 0$, it must be that the utility U offered by the entrant satisfies $U = u^j(\theta_I, p)$ because indifferent flexible consumers choose the entrant and offering higher U would reduce the entrant's profit. It then follows from (11) and (12) that, when less than one,

$$\widehat{\gamma}^{NP}(U,\theta_I,p) = 1 - \varphi\left(U,\alpha,\theta_I,p\right),$$

which is thus increasing in α for any U and, in particular, for $U = \tilde{U}_E^{FP}\left(\theta_E,p\right)$.

Next note that $\varphi(.)$ is decreasing in U when $U \geq U^*(\alpha,\theta_I,p)$ because the numerator is then the maximum value function of a problem in which an increase in U corresponds to a tighter constraint. To show $\widehat{\gamma}^{NP}(\widetilde{U}_E^{FP}(\theta_E,p),\theta_I,p)$ is then increasing in θ_E, it thus suffices to note that $\tilde{U}_E^{FP}\left(\theta_E,p\right)$ is increasing in θ_E. From the definition of $\widehat{\gamma}^j(U,\theta_I,p)$ in (12) and $\overline{U}^j\left(\gamma,\theta_I,p\right)$ in (13),

$$\gamma = \frac{v_I^{NP}\left(U^*\left(\alpha,\theta_I,p\right),\theta_I,p\right) - v_I^{NP}(\bar{U}^{NP}(\gamma,\theta_I,p),\theta_I,p)}{v_I^{NP}\left(U^*\left(\alpha,\theta_I,p\right),\theta_I,p\right)}.$$

The right-hand side of this is just $1-\phi\left(U,\alpha,\theta_I,p\right)$ evaluated at $U = \overline{U}^j\left(\gamma,\theta_I,p\right)$. It has already been shown that $1-\phi\left(U,\alpha,\theta_I,p\right)$ is increasing in α for any U and it was previously shown that $v_I^{NP}\left(U,\theta_I,p\right)$ is decreasing in U, which suffices to complete the result.

Proof of Proposition 5 From Proposition 2, rigid consumers receive utility $u^j\left(\theta_I,p\right)$ when the incumbent is type $j \in \{FP,NP\}$, which is exactly the same as when type j is a monopoly provider, so the result for them follows from Proposition 1. Also from Proposition 2, the result certainly holds for flexible consumers if the utility ranking claimed in the proposition holds. To establish that ranking, note that $\tilde{U}_I^j\left(\theta_I,p\right) \geq \bar{U}^j(\gamma,\theta_I,p)$ for $j \in \{FP,NP\}$ follows from the definition of $\tilde{U}_I^j\left(\theta_I,p\right)$ as the highest utility a type j incumbent with efficiency parameter θ_I can feasibly deliver. So, to establish the proposition, it remains to show only that $\bar{U}^{NP}(\gamma,\theta_I,p) > \tilde{U}_I^{FP}\left(\theta_I,p\right)$. Suppose contrary to this that $\bar{U}^{NP}(\gamma,\theta_I,p) \leq \tilde{U}_I^{FP}\left(\theta_I,p\right)$. Then, from (3),

$$\begin{aligned}
p\theta_I &= c\left(\tilde{U}_I^{FP}\left(\theta_I,p\right)\right) \\
&\quad > c\left(\tilde{U}_I^{FP}\left(\theta_I,p\right) - \hat{q}^*(\bar{U}^{NP}\left(\gamma,\theta_I,p\right),\theta_I,p)\right) + c\left(\hat{q}^*(\bar{U}^{NP}\left(\gamma,\theta_I,p\right),\theta_I,p)\right) \\
&\quad \geq c\left(\bar{U}^{NP}(\gamma,\theta_I,p) - \hat{q}^*(\bar{U}^{NP}\left(\gamma,\theta_I,p\right),\theta_I,p)\right) + c\left(\hat{q}^*(\bar{U}^{NP}\left(\gamma,\theta_I,p\right),\theta_I,p)\right) \\
&\quad = c\left(\hat{Q}^*\left(\bar{U}^{NP}\left(\gamma,\theta_I,p\right)\right),\theta_I,p\right) + c\left(\hat{q}^*\left(\bar{U}^{NP}\left(\gamma,\theta_I,p\right),\theta_I,p\right)\right) \\
&\quad = p\theta_I.
\end{aligned}$$

The second line of this follows because c is strictly convex, the third line from the supposition that $\bar{U}^{NP}(\gamma,\theta_I,p) \leq \tilde{U}_I^{FP}\left(\theta_I,p\right)$, the fourth line from (8) and the final line because the breakeven constraint for a not-for-profit incumbent (5) holds with equality.

But this gives a contradiction, so it must be that $\bar{U}^{NP}(\gamma, \theta_I, p) > \tilde{U}_I^{FP}(\theta_I, p)$.

Proof of Proposition 6 Consider first the case where the upper line in (15) in Proposition 2 applies when entry occurs and note that $u^k(\theta_E, p) \leq \overline{U}^{NP}(\gamma, \theta_I, p)$ implies $u^k(\theta_E, p) \leq \widetilde{U}_I^{NP}(\theta_I, p)$. The flexible consumers' payoff when entry occurs, evaluated with their decision preferences, is then $\tilde{U}_I^{NP}(\theta_I, p)$. The qualities delivering this are those that maximize consumers' decision preferences subject to the profit constraint (5), which are just those that a not-for-profit provider would choose if α were equal to 1, specifically $Q^*(1, \theta_I, p)$ and $q^*(1, \theta_I, p)$. Thus flexible consumers' true welfare given entry is $\alpha Q^*(1, \theta_I, p) + q^*(1, \theta_I, p)$. All we then need to show is that this is less than the true welfare of flexible consumers when there is no entry, which is just the payoff of the not-for-profit incumbent without entry. That is, we need to show

$$\alpha Q^*(\alpha, \theta_I, p) + q^*(\alpha, \theta_I, p) > \alpha Q^*(1, \theta_I, p) + q^*(1, \theta_I, p). \tag{A.10}$$

But $Q^*(\alpha, \theta_I, p)$ and $q^*(\alpha, \theta_I, p)$ are the unique maximizers of $\alpha Q + q$ subject to the profit constraint (5), so (A.10) certainly holds for any $\alpha > 1$.

Now consider the case where flexible consumers' payoff when entry occurs, evaluated with their decision preferences, is given by the lower line in (15) in Proposition 2. That is, it is given by $\overline{U}^{NP}(\gamma, \theta_I, p)$ that satisfies $\gamma = \widehat{\gamma}^{NP}(\overline{U}^{NP}(\gamma, \theta_I, p), \theta_I, p)$. Denote by $\widehat{Q}$ and $\widehat{q}$ the qualities that deliver this payoff. Then deterring entry is optimal if

$$\alpha Q^*(\alpha, \theta_I, p) + q^*(\alpha, \theta_I, p) > \alpha \widehat{Q} + \widehat{q}.$$

For exactly the same reasons as in the previous case, this always holds.

Proof of Proposition 7 From (17), welfare decreases with entry if

$$\frac{\bar{U}^{NP}(\gamma, \theta_I, p)}{U^*(\alpha, \theta_I, p)} < \frac{1 - (1 - \gamma)\nu}{\gamma}. \tag{A.11}$$

Moreover, $\overline{U}^{NP}(\gamma, \theta_I, p)$ defined by (13) is decreasing in γ and converges to $u^{NP}(\theta_I, p) = U^*(\alpha, \theta_I, p)$ as γ goes to zero. So $U^*(\alpha, \theta_I, p) / \bar{U}^{NP}(\gamma, \theta_I, p)$ is increasing in γ and converges to one as γ converges to zero. Thus, there exists $\hat{\gamma}$ such that $\gamma \leq U^*(\alpha, \theta_I, p) / \bar{U}^{NP}(\gamma, \theta_I, p)$ for all $\gamma \leq \hat{\gamma}$. Then, for all $\gamma \leq \hat{\gamma}$, there exists $\hat{\nu}$ such that

$$\frac{\bar{U}^{NP}(\gamma, \theta_I, p)}{U^*(\alpha, \theta_I, p)} = \frac{1 - (1 - \gamma)\hat{\nu}}{\gamma}$$

and, for all $\nu \leq \hat{\nu}$, (A.11) holds.

Proof of Proposition 8 First note that the welfare criterion (20) is differentiable with respect to p_E. So, for an interior solution $\hat{\theta}_E(p_E, p_I) \in (\underline{\theta}, \bar{\theta})$, any optimal p_E^* satisfies the first-order condition

$$-g(\hat{\theta}_E(p_E^*, p_I))[\Delta U(p_I) + \xi(p_I - p_E^*)]\frac{\partial\hat{\theta}_E(p_E^*, p_I)}{\partial p_E} - [1 - G(\hat{\theta}_E(p_E^*, p_I))]\xi = 0. \quad \text{(A.12)}$$

Moreover, from (19),

$$\frac{\partial\hat{\theta}_E(p_E, p_I)}{\partial p_E} = -\frac{\hat{\theta}_E(p_E, p_I)}{p_E}. \quad \text{(A.13)}$$

Use of (A.13) in the first-order condition (A.12) and re-arrangement gives (21).

Next note from (A.13) that

$$\frac{\partial^2\hat{\theta}_E(p_E, p_I)}{\partial p_E^2} = -\frac{1}{p_E}\left[p_E\frac{\partial\hat{\theta}_E(p_E, p_I)}{\partial p_E} - \hat{\theta}_E(p_E, p_I)\right] > 0. \quad \text{(A.14)}$$

Moreover, the derivative of the left-hand side of (21), with (A.13) substituted for $\hat{\theta}_E(p_E^*, p_I)$, with respect to p_E can be written

$$-\left[\frac{\Delta U(p_I)}{\xi} + (p_I - p_E^*)\right]\frac{\partial^2\hat{\theta}_E(p_E^*, p_I)}{\partial p_E^{*2}} + \frac{\partial\hat{\theta}_E(p_E^*, p_I)}{\partial p_E^*},$$

which is negative given (A.13), (A.14) and that the square bracket on the left-hand side of (21) must be positive for any p_E^* that satisfies (21) with $\hat{\theta}_E(p_E, p_I) \in (\underline{\theta}, \bar{\theta})$. Furthermore, by Corollary 2 in Bagnoli and Bergstrom (2005), $g(\theta_E)$ log-concave implies that $[1 - G(\theta_E)]/g(\theta_E)$ is monotone decreasing in θ_E. Since $\hat{\theta}_E(p_E, p_I)$ is decreasing in p_E, the right-hand side of (21) must, therefore, be increasing in p_E^*. Thus, there cannot be more than one value of p_E^* that satisfies (21).

Proof of Proposition 9 Compare the expression for welfare in (24) with competition ($\hat{\theta}_E < \bar{\theta}$) and without competition ($\hat{\theta}_E = \bar{\theta}$) to yield the welfare gain from allowing entry

$$-\gamma\delta_I\left[1 - G\left(\hat{\theta}_E\right)\right]v_I^j\left(u^j(\theta_I, p), \theta_I, p\right)$$
$$+\gamma\int_{\hat{\theta}_E}^{\bar{\theta}}\left[\delta_E v_I^j(\hat{U}(\theta_I, \gamma, p), \theta_I, p) + \hat{U}(\theta_I, \gamma, p) - u^j(\theta_I, p)\right]dG(\theta). \quad \text{(A.15)}$$

This is negative for δ_I sufficiently large.

Appendix B A General Formulation

B.1 General Model and Results

This appendix provides a more general formulation of the core ideas where, instead of having only two groups of consumers, we allow for the possibility that any consumer is willing to switch to the entrant. We also allow for more than two dimensions of quality. The main aim of the section is to show that the core insights from the model in the main text carry over to this more general setting.

Suppose then that consumers differ in their benefit $b \in \left[\underline{b}, \overline{b}\right]$ from switching to the entrant, with distribution function $F(b)$ that admits a density and is log-concave.[18] We make no assumption about the signs of $\underline{b}$ and $\overline{b}$, so consumers may prefer to stay with the incumbent, or to switch to the entrant, when offered the same quality levels by both.

The continuous benefit from switching generalizes the idea of rigid and flexible consumers. This benefit can arise for a variety of reasons that are relevant for schools and hospitals, reflecting, for example, the geographical location of the incumbent or entrant which makes use of one of the providers more convenient for some consumers. It might also proxy for other intrinsic attributes.

Our general formulation also allows for vectors of both types of quality. Specifically, let q^1 be an M-element vector of observable qualities, with generic element q_m^1, that a provider can commit to before consumers choose their provider, q^2 be an N-element vector of unobservable qualities, with generic element q_n^2, to which commitment is infeasible before consumers choose their provider, and q be the overall vector of qualities $\left(q^1, q^2\right)$. For notational convenience let π denote the vector of parameters in the model.[19] All consumers have the same utility $U(q, \pi)$ from provision by the incumbent, which is everywhere strictly increasing in each element of q. A consumer with switching benefit b has utility $U(q, \pi) + b$ from being served by an entrant that provides quality vector q. As before, consumers choose provision if and only if they attain utility of at least zero and those indifferent between providers choose the entrant.

Providers have constant returns to scale and serve all consumers who come to them.[20] They enter the market if and only if they achieve a positive payoff from doing so and the order of moves is the same as in the main text.

As before, for $j \in \{FP, NP\}$ and $i \in \{I, E\}$, let $u^j(\pi)$ be the utility to consumers delivered by a type j provider if not constrained by competition and $\widetilde{U}_i^j(\pi) > u^j(\pi)$ be

[18] This is weaker than the more widely used assumption that the density F' is log-concave; see Jewitt (1987) for discussion.

[19] For the model in the main text, $\pi = (\theta_I, \theta_E, p_I, p_E, \alpha, \lambda, \beta)$. However, the parameterization in the generalized model can be richer than that.

[20] It is straightforward to introduce an entry cost.

the highest consumer payoff type j is willing to provide, but now both net of switching benefit. Also, let $v_i^j(U,\pi)$ be the highest payoff available to provider type j if delivering consumer utility U conditional on having entered the market. This is assumed to be continuously differentiable and strictly concave in U for all U in excess of what the provider would offer in the absence of competition.[21]

Conditional on utility offers U_I and U_E from the incumbent and entrant respectively, both net of switching benefit, consumers with switching benefit b choose I if $U_I > U_E + b$. Let

$$b^*(U_I, U_E) = \begin{cases} \underline{b}\,, & \text{if} \quad U_I < U_E + \underline{b}; \\ U_I - U_E\,, & \text{if} \quad U_E + \underline{b} \le U_I \le U_E + \overline{b}; \\ \overline{b}\,, & \text{if} \quad U_I > U_E + \overline{b}; \end{cases} \tag{B.1}$$

be the value of b that determines consumer choices given U_I and U_E and let $U_I^j(U_E,\pi)$ denote the best response utility offer for a type $j \in \{FP, NP\}$ incumbent if the entrant offers U_E. We assume $\underline{b}$ sufficiently low that the incumbent always chooses to retain some consumers. For this generalized formulation, the following result corresponds to Proposition 2.

Proposition 10 *For $j,k \in \{FP, NP\}$, a sufficient condition for entry by a type k potential entrant facing a type j incumbent is that $\widetilde{U}_I^j(\pi) < \widetilde{U}_E^k(\pi) + \overline{b}$. For $\widetilde{U}_I^j(\pi) \ge \widetilde{U}_E^k(\pi) + \overline{b}$, a necessary and sufficient condition for entry by a type k potential entrant facing a type j incumbent is*

$$\frac{\partial}{\partial U_I} v_I^j(U_I,\pi) + v_I^j(U_I,\pi)\,F'\left(\overline{b}\right) \le 0, \qquad \textit{for} \quad U_I = \widetilde{U}_E^k(\pi) + \overline{b}, \qquad j,k \in \{FP, NP\}. \tag{B.2}$$

If the incumbent would set $u^j(\pi) \in \left(0, \widetilde{U}_I^i(\pi)\right)$ in the absence of entry and entry occurs, all consumers strictly gain from entry.

Entry occurs as long as the entrant has a non-negative payoff from servicing the consumers with the highest benefit from switching, those with $b = \bar{b}$. Thus there is entry for sure if the highest utility the entrant is willing to offer attracts some consumers even when the incumbent also offers the highest utility it is willing to offer (that is, if $\widetilde{U}_I^j(\pi) < \widetilde{U}_E^k(\pi) + \overline{b}$). Otherwise, there is entry if and only if the incumbent prefers to cede part of the market at the highest utility the entrant is willing to offer, a condition captured by (B.2), which generalizes (14).

The main economic difference from this more general formulation is that even consumers who do not switch to an entrant can strictly gain from entry,[22] which strength-

[21] These properties are satisfied by the specific functional forms in the main text.

[22] That is certainly the case if the incumbent is a not-for-profit that offers strictly positive utility in the absence of entry (that is, $u^{NP}(\pi) > 0$) because then, with all consumers potentially flexible, it is always worth the incumbent offering at least marginally higher utility to retain some additional consumers.

ens the welfare results. This is because competition may lead the incumbent to offer higher utility to retain additional consumers.

For the model of the main text, the probability of entry by a for-profit provider is lower with a not-for-profit incumbent than with a for-profit incumbent. The next result gives a general condition for any parameter change to reduce the probability of entry in the generalized model.

Proposition 11 *For $j,k \in \{FP, NP\}$, consider an equilibrium that, conditional on entry, has U_E such that $U_I^j(U_E,\pi) \in (0, \tilde{U}_I^j(\pi))$. A change in any parameter in π that increases $\frac{\partial v_I^j\left(\widetilde{U}_E^k(\pi)+\bar{b},\pi\right)/\partial U}{v_I^j\left(\widetilde{U}_E^k(\pi)+\bar{b},\pi\right)}$ but does not affect $v_E^k(U_E,\pi)$ reduces the probability of entry.*

This proposition shows that the finding that entry is less likely with a not-for-profit incumbent than with a for-profit one extends beyond the particular formulation in the main text. There are two potential channels at work here. The first is a cost channel; a not-for-profit incumbent that provides a positive (instead of a zero) level of some unobserved quality can deliver given utility at lower cost even with the same (strictly convex) cost function. That results in an increase in optimal U_I for given U_E. With $v_E^k(U_E,\pi)$ unaffected, this increases the critical value of θ_E at which entry becomes worthwhile and hence, for a given distribution of θ_E, reduces the probability of entry. The second is a payoff channel which depends on how a change in parameter that affects preferences changes the incentive of an incumbent to offer a particular level of U.[23]

It is also instructive to see how the result on encouraging or discouraging entry in Proposition 8 is changed in the more general formulation of this section. To generalize

It may not be the case with a for-profit incumbent who, as in the previous model, sets $u^{FP}(\pi) = 0$. Then the incumbent may prefer to offer $U_I^{FP}(U_E,\pi) = 0$ for some U_E even with entry and serve only those consumers with highly negative switching benefits if the distribution F is such that there are sufficient of these. Formally, the difference between $u^j(\pi) > 0$ and $u^j(\pi) = 0$ is that the former is an interior solution at which a marginally higher utility always attracts more consumers when entry occurs, whereas the latter is a corner solution.

[23]The following example illustrates the payoff channel at work. Suppose the not-for-profit incumbent has payoff function $\alpha U(q,\pi) + \Pi(q,\pi)$, where $\Pi(q,\pi)$ is its profit and $\alpha > 0$, and let $q(U,\pi)$ denote the incumbent's optimal choice of quality vector to deliver utility U given the constraints it faces. Then

$$v_I^j(U,\pi) = \alpha U(q(U,\pi)) + \Pi(q(U,\pi),\pi).$$

This reduces to the for-profit payoff function $v_I^{FP}(U,\pi)$ for $\alpha = 0$. Since α affects profit only through the choice of q, it follows from the envelope theorem that $\partial v_I^j(U,\pi)/\partial\alpha = U(q(U,\pi)) > 0$, and hence $\partial^2 v_I^j(U,\pi)/\partial U\partial\alpha = 1$, even if there were no change in unobservable qualities. Moreover, for any best response, $v_I^j(U,\pi)$ is non-increasing in U. Straightforward differentiation then establishes that $\frac{\partial v_I^j(U,\pi)/\partial U}{v_I^j(U,\pi)}$ is increasing in α as long as $U < \tilde{U}_I^j(\pi)$. That also results in an increase in optimal U_I for given U_E. With $v_E^k(U_E,\pi)$ unaffected, this increases the critical efficiency level at which entry becomes worthwhile. Thus an increase in α from zero (which corresponds to moving from a for-profit incumbent to a not-for-profit incumbent) reduces the probability of entry.

the welfare criterion in (20), it is helpful to define the parameter vector $\hat{\pi}$ as the parameter vector π excluding the efficiency parameter of the potential entrant θ_E and the payment to the entrant p_E. (That is, $\hat{\pi} = \pi \backslash (\theta_E, p_E)$.) For consistency with the earlier model, the entrant's cost of supplying quality is decreasing in θ_E. Then, with $\hat{}$ used to specify equilibrium values conditional on the parameters, the welfare criterion given incumbent type j and potential entrant type k is

$$G\left(\hat{\theta}_E\left(p_E,\hat{\pi}\right)\right) u^j\left(\hat{\pi}\right) - \xi p_I + \int_{\hat{\theta}_E(p_E,\hat{\pi})}^{\bar{\theta}} \left\{ F\left(\hat{b}\left(\theta_E,p_E,\hat{\pi}\right)\right) \hat{U}_I^j\left(\theta_E,p_E,\hat{\pi}\right) \right.$$
$$\left. + \int_{\hat{b}(\theta_E,p_E,\hat{\pi})}^{\bar{b}} \left[\hat{U}_E^k\left(\theta_E,p_E,\hat{\pi}\right) + b - \xi\left(p_E - p_I\right)\right] dF\left(b\right) \right\} dG\left(\theta_E\right), \quad \text{(B.3)}$$

where $\hat{\theta}_E(p_E,\hat{\pi})$ denotes the entrant efficiency at which entry becomes just worthwhile and $\hat{b}\left(\theta_E,p_E,\hat{\pi}\right) = \hat{U}_I^j\left(\left(\theta_E,p_E,\hat{\pi}\right)\right) - \hat{U}_E^k\left(\theta_E,p_E,\hat{\pi}\right)$. The following result is the counterpart to Proposition 8.

Proposition 12 *Suppose, for $j,k \in \{FP,NP\}$, a type j incumbent competes with a type k potential entrant and $\hat{\theta}_E(p_E,\hat{\pi}) < \bar{\theta}$ at $p_E = p_I$. Then a policy-maker increases welfare by encouraging entry by increasing p_E above p_I if*

$$-\left[\hat{U}_E^k\left(\hat{\theta}_E\left(p_E,\hat{\pi}\right),p_E,\hat{\pi}\right) + F\left(\hat{b}\left(\hat{\theta}_E\left(p_E,\hat{\pi}\right),p_E,\hat{\pi}\right)\right)\hat{b}\left(\hat{\theta}_E\left(p_E,\hat{\pi}\right),p_E,\hat{\pi}\right)\right.$$
$$\left. + \int_{\hat{b}\left(\hat{\theta}_E(p_E,\hat{\pi}),p_E,\hat{\pi}\right)}^{\bar{b}} bF'\left(b\right)db - u^j\left(\hat{\pi}\right)\right] g\left(\hat{\theta}_E\left(p_E,\hat{\pi}\right)\right)\frac{\partial\hat{\theta}_E\left(p_E,\hat{\pi}\right)}{\partial p_E}$$
$$+ \int_{\hat{\theta}_E(p_E,\hat{\pi})}^{\bar{\theta}} \left\{ \frac{\partial\hat{U}_E^k\left(\theta_E,p_E,\hat{\pi}\right)}{\partial p_E} + F\left(\hat{b}\left(\theta_E,p_E,\hat{\pi}\right)\right)\frac{\partial\hat{b}\left(\theta_E,p_E,\hat{\pi}\right)}{\partial p_E}\right.$$
$$\left. - \left[1 - F\left(\hat{b}\left(\theta_E,p_E,\hat{\pi}\right)\right)\right]\xi \right\} dG\left(\theta_E\right) > 0 \quad \text{(B.4)}$$

and discouraging it if the strict inequality is reversed.

The term in square brackets on the top two lines of (B.4) is the utility gain to those consumers who would have switched to an entrant with cost parameter $\hat{\theta}$ from having entry occur at a marginally lower cost parameter as the result of the marginal increase in the payment to the entrant. Unlike in the simple model, it involves an integral term because those consumers switching to the entrant differ in their benefit from doing so. The lower two lines of (B.4) incorporate the welfare effect of the *change* in the proportion of consumers who switch to the entrant because the payment to the entrant affects the utility the entrant offers those who switch. This second effect does not arise in the simple model because there the proportion of consumers who switch is fixed. This second effect complicates evaluation of having different payments for the entrant and the incumbent. But the essential point, in line with the simpler model above, is that there is no more reason to presume that it is optimal to set the same payment for both incumbent and entrant when all consumers are potentially flexible than when only a fixed proportion are.

Overall, the results in this appendix confirm that a range of insights generated by the simple model are indeed robust to having a continuous benefit from switching and arbitrary dimensions of quality. It should also be clear that we do not need to stick to the specific way that we modeled not-for-profit preferences for the core results to hold as long as they satisfy the key assumptions outlined here.[24]

B.2 Proofs for General Model

Lemma 1 *A type j incumbent's best response to an entrant offering U_E that attracts some consumers is the unique $U_I^j(U_E,\pi)$ that satisfies*

$$F\left(b^*\left(U_I^j(U_E,\pi),U_E\right)\right)\frac{\partial}{\partial U_I}v_I^j\left(U_I^j(U_E,\pi),\pi\right)$$
$$+\,v_I^j\left(U_I^j(U_E,\pi),\pi\right)F'\left(b^*\left(U_I^j(U_E,\pi),U_E\right)\right)=0 \quad \text{(B.5)}$$

or, equivalently,

$$-\frac{\partial v_I^j\left(U_I^j(U_E,\pi),\pi\right)/\partial U_I}{v_I^j\left(U_I^j(U_E,\pi),\pi\right)}=\frac{F'\left(b^*\left(U_I^j(U_E,\pi),U_E\right)\right)}{F\left(b^*\left(U_I^j(U_E,\pi),U_E\right)\right)}. \quad \text{(B.6)}$$

Proof. A type j incumbent's best response to an entrant offering U_E is $U_I^j(U_E,\pi)$ that satisfies

$$U_I^j(U_E,\pi)\in\arg\max_{U_I}v_I^j(U_I,\pi)\,F\left(b^*(U_I,U_E)\right). \quad \text{(B.7)}$$

The first-order necessary condition for this best response to be interior (that is, with $U_I\in\left(0,\widetilde{U}_I^i(\pi)\right)$ such that $b^*(U_I,U_E)\in(\underline{b},\bar{b})$) is (B.5) because, from (B.1), $\partial b^*(U_I,U_E)/\partial U_I=1$ for $b^*(U_I,U_E)\in(\underline{b},\bar{b})$. Moreover, (B.5) can be written as (B.6). With $v_I^j(U,\pi)$ non-negative and strictly concave in U in the relevant range, the left-hand side of (B.6) is strictly increasing in U_I. With F log concave, F'/F is non-increasing, so the right-hand side of (B.6) is non-increasing in U_I since $\partial b^*(U_I,U_E)/\partial U_I=1$ at any interior solution from (B.1). There can, therefore, be at most one solution to (B.5) with $U_I\in\left(0,\widetilde{U}_I^i(\pi)\right)$ such that $b^*(U_I,U_E)\in(\underline{b},\bar{b})$ and hence, by continuity, at most one $U_I^j(U_E,\pi)$ that satisfies (B.7) with $b^*(U_I,U_E)\in(\underline{b},\bar{b})$. By assumption, $\underline{b}$ is sufficiently low that the

[24] In Section B.3 of this appendix, we give a specific parameterized example where all of these assumptions are satisfied.

incumbent always chooses to retain some consumers, so U_I such that $b^*(U_I, U_E) = \underline{b}$ cannot be a best response and $b^*(U_I, U_E) = \bar{b}$ corresponds to no entry. Thus there is at most one $U_I^j(U_E, \pi)$ that satisfies (B.7) for given U_E at which entry can occur and this satisfies (B.5) and (B.6). ■

Proof of Proposition 10 With constant returns to scale, the potential entrant enters if and only if it can attract at least the consumers with the largest benefit from switching $\bar{b}$. The proof considers separately the sufficient conditions for entry, the necessary condition for entry, and consumer utility conditional on entry as specified in the proposition.

Sufficient conditions for entry: If $\widetilde{U}_I^j(\pi) < \widetilde{U}_E^k(\pi) + \bar{b}$, the potential entrant is prepared to offer a higher payoff to type $\bar{b}$ consumers than the incumbent is prepared to offer them, so entry is worthwhile. For $\widetilde{U}_I^j(\pi) \geq \widetilde{U}_E^k(\pi) + \bar{b}$, suppose (B.2) holds. From Lemma 1, there is at most one solution to (B.5) so the incumbent would not increase its payoff by offering more than $\widetilde{U}_E^k(\pi) + \bar{b}$ to retain the consumers with the greatest benefit from switching to the entrant. The entrant would be prepared to offer $\widetilde{U}_E^k(\pi)$ to attract those consumers.

Necessary condition for entry: Suppose (B.2) does not hold. Then, even if the entrant offers the highest consumer payoff it is prepared to offer to attract the consumers, $\widetilde{U}_E^k(\pi)$, the incumbent's payoff is increasing in U_I at the value that retains even the consumers with the greatest benefit from switching $\bar{b}$. Moreover, with at most one solution to (B.5), the incumbent would obtain a lower payoff by offering any lower U_I.

Consumer utility conditional on entry: In the absence of entry, a type j incumbent chooses U_I to satisfy (B.7) given $b^*(U_I, U_E) = \bar{b}$ so $F(b^*(U_I, U_E)) = 1$ for all $U_I \geq 0$. By definition, the solution to that is $u^j(\pi)$, the payoff to all consumers in the absence of entry. If $u^j(\pi) > 0$, it must satisfy $\partial v_I^j(u^j(\pi), \pi) / \partial U_I = 0$. Conditional on entry, the part of the left-hand side of (B.5) on the lower line is strictly positive for $U_I = u^j(\pi) < \widetilde{U}_I^i(\pi)$, as assumed. With $v_I^j(U, \pi)$ strictly concave in U, that implies $U_I^j(U_E, \pi) > u^j(\pi)$. Thus even consumers who do not switch to the entrant receive strictly higher utility conditional on entry as, *a fortiori*, do those who choose to switch to the entrant.

Proof of Proposition 11 By assumption, $\underline{b}$ is sufficiently low that the incumbent always chooses to retain some consumers, so U_I such that $b^*(U_I, U_E) = \underline{b}$ cannot be a best response and $b^*(U_I, U_E) = \bar{b}$ corresponds to no entry. Thus, conditional on entry, the incumbent's best response $U_I^j(U_E, \pi)$ to U_E such that $U_I^j(U_E, \pi) \in \left(0, \widetilde{U}_I^j(\pi)\right)$ is, from Lemma 1, given by the unique solution to (B.6). A change in any parameter in π that increases $\frac{\partial v_I^j(U,\pi)/\partial U}{v_I^j(U,\pi)}$ for all $U \in (0, \tilde{U}_I^j(\pi))$ reduces the left-hand side of (B.6) for all $U \in (0, \tilde{U}_I^j(\pi))$, which implies an increase in optimal U_I for given U_E. But an increase in optimal U_I for given U_E with $v_E^k(U_E, \pi)$ unaffected increases the critical value of θ_E at which entry becomes worthwhile and hence, for a given distribution of θ_E, reduces the probability of entry.

Proof of Proposition 12 Substitution for $\hat{U}_I^j(\theta_E, p_E, \hat{\pi})$ in (B.3) using $\hat{b}(\theta_E, p_E, \hat{\pi}) = \hat{U}_E^k(\theta_E, p_E, \hat{\pi}) - \hat{U}_I^j(\theta_E, p_E, \hat{\pi})$ and differentiation with respect to p_E, with p_E set equal to p_I, yields the left-hand side of (B.3). If this is strictly positive, welfare is increased by raising p_E above p_I. If it is strictly negative, welfare is increased by reducing p_E below p_I, as claimed in the proposition.

B.3 Example with multiple qualities

The following example with multiple qualities exhibits properties of $v_i^j(U, \pi)$ that satisfy the assumptions in Appendix B.1. Suppose that the not-for-profit provider has objective function

$$\alpha U(q, \pi) + \Pi(q, \pi), \tag{B.8}$$

where $\Pi(q, \pi)$ is its profit function and $\alpha > 0$, and the utility and profit functions have the forms

$$U(q, \pi) = \sum_{n=1}^{N} r_n q_n; \; \Pi(q, \pi) = p_i - \frac{1}{2\theta_i} \sum_{n=1}^{N} q_n^2, \quad \text{with} \quad \theta_i > 0, \;\text{for} \quad i \in \{I, E\}, \tag{B.9}$$

with $r_n > 0$ for all $n = 1, \ldots, N$, normalized so that $\sum_{m=1}^{M} r_m = 1$, and the other notation as before. We can think of q_n as the square root of the relative monetary expenditure on quality dimension n and r_n as the linearized marginal utility of additional q_n at the enforceable level of quality 0.

A provider chooses quality dimensions $n = M+1, \ldots, N$ to maximize its payoff subject only to the breakeven constraint and non-negativity of the q_n because consumers have already chosen their provider. Its optimization problem at this stage is

$$\max_{q_n, n=M+1,\ldots,N} \left\{ \alpha \sum_{n=M+1}^{N} r_n q_n - \frac{1}{2\theta_i} \sum_{n=M+1}^{N} q_n^2 \right\} \text{ subject to} \tag{B.10}$$

$$p_i - \frac{1}{2\theta_i} \sum_{n=1}^{N} q_n^2 \geq 0 \tag{B.11}$$

$$q_n \geq 0, \text{ for } n = M+1, \ldots, N, \text{ and given } q_m, \text{ for } m = 1, \ldots, M. \tag{B.12}$$

The first-order condition for an interior solution to q_n is

$$\alpha r_n - \frac{1 + \lambda_i}{\theta_i} q_n = 0, \quad \text{for } n = M+1, \ldots, N,$$

where $\lambda_i \geq 0$ is a multiplier satisfying a complementary inequality with the breakeven constraint (B.11). This gives the solution

$$q_n = r_n \frac{\alpha\theta_i}{1+\lambda_i}, \text{ for } n = M+1, \ldots, N; i \in \{I, E\}. \tag{B.13}$$

Provider i's optimization problem for quality dimensions $m = 1, \ldots, M$ (chosen before consumers have chosen a provider) to deliver utility U must ensure that the q_n satisfy (B.13) and is thus

$$\max_{q_m, m=1,\ldots,M} \left\{ \alpha \sum_{m=1}^{N} r_m q_m + [p_i - \frac{1}{2\theta_i} \sum_{m=1}^{N} q_m^2] \right\} \text{ subject to} \tag{B.14}$$

$$p_i - \frac{1}{2\theta_i} \sum_{m=1}^{N} q_m^2 \geq 0 \tag{B.15}$$

$$\sum_{m=1}^{N} r_m q_m \geq U \tag{B.16}$$

$$q_m \geq 0, \text{ for } m = 1, \ldots, M, \text{ and } q_n, \text{ for } n = M+1, \ldots, N, \text{ satisfies (B.13).} \tag{B.17}$$

If U is sufficiently high that the breakeven constraint is binding, $\Pi(q, \pi) = 0$, so a provider with $\alpha > 0$ must be maximizing $\alpha U(q)$, or equivalently $U(q)$, subject to the breakeven constraint regardless of the specific value of α (as long as it is strictly positive). This corresponds to delivering the highest utility that is feasible given the constraints that, for the purposes of this example, we denote $\tilde{U}_i^{\alpha}(\pi)$. Moreover, a for-profit provider with $\alpha = 0$ delivers $\tilde{U}_i^{0}(\pi)$ if its profits are zero. Thus, if $U < \tilde{U}_i^{\alpha}(\pi)$, the breakeven constraint is not binding. Then λ_i in (B.13) is zero and the first-order condition for an interior solution to q_m is

$$r_m(\alpha + \mu_i) - \frac{1}{\theta_i} q_m = 0, \text{ for } m = 1, \ldots, M,$$

where $\mu_i \geq 0$ is a multiplier satisfying a complementary inequality with the utility constraint (B.16). So

$$q_m = r_m (\alpha + \mu_i) \theta_i, \text{ for } m = 1, \ldots, M. \tag{B.18}$$

Use of (B.13) with $\lambda = 0$ and (B.18) in the utility function in (B.9), along with the normalization $\sum_{m=1}^{M} r_m^2 = 1$ and $R = \sum_{n=M+1}^{N} r_n^2$, gives utility

$$(\alpha + \mu_i) \theta_i + \alpha\theta_i R.$$

If this satisfies the utility constraint with $\mu_i = 0$, q_m is given by (B.18) with $\mu_i = 0$. If not, then μ_i must satisfy

$$(\alpha + \mu_i)\theta_i = U - \alpha\theta_i R.$$

Used in (B.18), these give

$$q_m = r_m \max\left\{\alpha\theta_i, (U - \alpha\theta_i R)\right\}, \quad \text{for } U \in \left[0, \tilde{U}_i^\alpha(\pi)\right), \ m = 1, \ldots, M. \tag{B.19}$$

This and (B.13) for $\lambda = 0$ can be used in the profit function in (B.9) to give, when the breakeven constraint is not binding,

$$\Pi(q, \pi) = p_i - \frac{1}{2\theta_i} \sum_{m=1}^{M} \left(r_m \max\left\{\alpha\theta_i, U - \alpha\theta_i R\right\}\right)^2 - \frac{1}{2\theta_i} \sum_{n=M+1}^{N} \left(r_n \alpha\theta_i\right)^2$$

or, with $\sum_{m=1}^{M} r_m^2 = 1$ and $R = \sum_{n=M+1}^{N} r_n^2$,

$$\Pi(q, \pi) = p_i - \frac{1}{2\theta_i}\left[\left(\max\left\{\alpha\theta_i, U - \alpha\theta_i R\right\}\right)^2 + \left(\alpha\theta_i\right)^2 R\right]. \tag{B.20}$$

This can be used to check the conditions under which the breakeven constraint (B.11) is not binding. Specifically, the breakeven constraint is not binding if

$$\left(\max\left\{\alpha\theta_i, U - \alpha\theta_i R\right\}\right)^2 \leq 2\theta_i p_i - \left(\alpha\theta_i\right)^2 R. \tag{B.21}$$

$\tilde{U}_i^\alpha(\pi)$ satisfies (B.21) with equality.

Use of (B.20) in (B.8) gives the payoff to type α from delivering utility $U \in \left[0, \tilde{U}_i^\alpha(\pi)\right)$ as

$$v_i^\alpha(U, \pi) = \alpha U + p_i - \frac{1}{2\theta_i}\left[\left(\max\left\{\alpha\theta_i, U - \alpha\theta_i R\right\}\right)^2 + \left(\alpha\theta_i\right)^2 R\right], \qquad \text{for} \quad U \in \left[0, \tilde{U}_i^\alpha(\pi)\right).$$

Note that this function is identical to the case in which there is just one of each type of quality, with marginal utilities of 1 and R respectively. Moreover,

$$\frac{\partial v_i^\alpha(U, \pi)}{\partial U} = \begin{cases} \alpha, & \text{for} \quad U \in \left[0, \alpha\theta_i(1+R)\right), \\ \alpha(1+R) - \frac{U}{\theta_i}, & \text{for} \quad U \in \left[\alpha\theta_i(1+R), \tilde{U}_i^\alpha(\pi)\right). \end{cases}$$

This is positive for

$$U < \alpha\theta_i(1+R),$$

which implies that the utility $u^\alpha(\pi)$ offered by the incumbent in the absence of entry is $u^\alpha(\pi) = \max\{0, \alpha\theta_I(1+R)\}$, and it is continuous for $U \in \left[u^\alpha(\pi), \tilde{U}_i^\alpha(\pi)\right)$. It is always negative for $\alpha = 0$, in which case the utility constraint is always binding. Moreover, $\partial^2 v_i^\alpha(U, \pi)/\partial U^2 < 0$ for $U \in \left[\alpha\theta_i(1+R), \tilde{U}_i^\alpha(\pi)\right)$, so $v_i^\alpha(U, \pi)$ is strictly concave in U for $U \in \left[u^\alpha(\pi), \tilde{U}_i^\alpha(\pi)\right)$.