Theoretical Economics 4 (2009), 89–125 1555-7561/20090089

Monopolistic group design with peer effects

Simon Board
Department of Economics, University of California Los Angeles

In a range of settings, private firms manage peer effects by sorting agents into different groups, be they schools, communities, or product categories. This paper considers such a firm, which controls group entry by setting a series of anonymous prices. We show that private provision systematically leads to two distortions relative to the efficient solution: first, agents are segregated too finely; second, too many agents are excluded from all groups. We demonstrate that these distortions are a consequence of anonymous pricing and do not depend upon the nature of the peer effects. This general approach also allows us to assess the way the 'returns to scale' of peer technology and the cost of group formation affect the optimal group structure.

Keywords. Mechanism design, peer effects, public goods, network effects.
JEL classification. D82, H40, L12.

1. Introduction

In an increasingly privatized world, for-profit organizations have come to play an important role in many markets where peer effects are prominent. This paper considers such a market, where a firm posts a series of prices and agents self-select into different groups. The quality of a group, in turn, depends on the number and characteristics of its members. We show that private provision systematically leads to two distortions in group formation relative to the efficient solution. First, there is too much segregation between different types of agents; that is, groups are excessively homogenous. Second, too many agents are excluded from all available groups.

The model captures the key features of several important markets. First, consider the market for education, where peer effects play an important role in shaping students' goals and learning experiences. In this type of market, firms can manage peer effects to their advantage by charging more for courses and at institutions that attract above-average students. This type of differentiation is commonplace among providers of higher education and professional training and, with the introduction of vouchers, promises to become important among primary and secondary schools.

Simon Board: sboard@econ.ucla.edu
I have received helpful comments from Richard Arnott, Heski Bar-Isaac, Jeremy Bulow, Ed Lazear, John McMillan, Rob McMillan, Moritz Meyer-ter-Vehn, Andreas Park, Will Strange, Bob Wilson, Bill Zame, and Jeff Zwiebel. I thank seminar audiences at Brock, McMaster, Stanford, Toronto, York, IIOC 2006, and ESEM 2006. I am also grateful to the editor, Jeff Ely, and three referees who provided many insightful comments.

Second, the model can be used to study (indirect) network goods, such as cars, electronics, and games consoles. When buying one of these goods the consumer cares about the availability of post-purchase services, such as car repairs, accessories and software. The quality of these services, in turn, depends on the number and composition of consumers buying the good. A firm can therefore discriminate between different types of consumers by offering a range of different products.[1]

The third application concerns the market for private communities (e.g. condominiums, planned unit developments) which now house more than forty million Americans. When purchasing a unit, buyers care about the type of neighbours both directly (e.g. social ties, crime) and through the resulting neighbourhood services (e.g. shops, schools). A developer will then seek to design tiers of communities to attract different types of agents.

Finally, peer effects play an important role in the theory of the firm. Agents care about the composition of their division or team, through direct interactions or shared bonuses. Agents also care about the size of their division or team, since this may yield returns to scale or exacerbate moral hazard. The firm thus seeks to assemble compatible agents in order to maximize its productivity and minimize its wage bill.

A significant problem in analysing peer effects is that peer technology can be very complicated, encompassing composition effects, network effects, and congestion effects. Peer technology also differs greatly across environments. In a recent survey on the role of private education, Helen Ladd (2002, p. 14) writes

> "This lack of clarity about how peer effects differ among groups rules out any clear predictions about whether a voucher program would be likely to increase or decrease the overall productivity of the education system through the mechanism of peer effects".

Despite this concern, we analyze the distortions induced by private provision while placing very little structure on the nature of peer effects. This general approach enables us to examine how the degree of segregation depends on the form of peer effects. It also helps us interpret the recent empirical literature quantifying peer effects in different environments.

1.1 *Outline of the paper*

The basic structure of the model is as follows. First, a single principal posts a range of anonymous group-entry prices. Agents vary in their willingness to pay for group quality and, after observing these prices, sort themselves into different groups. The quality of a group, in turn, is a one-dimensional metric that depends upon the characteristics of its members. This quality function is allowed to be very general and subsumes the average

[1] With some goods, the consumer directly cares about the identity of the other purchasers: restaurants, golf clubs, and luxury good manufacturers all seek to affect the attractiveness of their product through exclusivity. For example, Kaneff owns six golf courses in Ontario, charging a range of fees and separating different types of customers into groups. Similarly, Ryanair has unassigned seating on its planes, but allows consumers to buy a priority boarding pass. See Rayo (2003) for other examples.

quality model (e.g. Rayo 2003), the Cobb–Douglas quality model (e.g. Epple and Romano 1998) and the multiplicative quality model (e.g. Lazear 2001).

Since pricing is anonymous, the principal must rely on agents to self-select into different groups. Self-selection immediately implies that agents who care more about group quality must be in better quality groups (the monotonicity condition). As a result, if the agents who generate high quality have a low willingness to pay, then the principal must assign all agents to groups of identical quality. Conversely, if the agents who generate high quality have a high willingness to pay, then the principal can segregate the agents into groups of different standards.

The paper first analyzes the principal's problem when group formation is costless, showing that profit-maximization leads to two distortions relative to the welfare-maximizing group structure. The first distortion, the *segregation effect*, states that there are too many groups under profit-maximization. Intuitively, by splitting a group into two, putting all the high types into one group and the low types into another, the principal increases the price the high types are willing to pay in order to avoid the low quality group. Crucially, we do not require any assumptions on the nature of peer effects in order to attain this result: the necessary restrictions come endogenously from the requirement that agents self-select into groups. This segregation effect implies that, if the optimal group structures are ordered, then the distribution of group qualities under private provision has a lower mean and tends to be more dispersed than the efficient distribution.

The second distortion, the *exclusion effect*, states that too many agents are excluded from all privately provided groups. The exclusion effect is analogous to the standard result that a monopolist prices above marginal cost. Intuitively, excluding some low types of agents raises the price paid by those agents who are not excluded.

We further analyze how the optimal group structure depends upon the nature of peer interactions. When a quality function exhibits negative returns to scale, in that merging two groups lowers the average quality, then welfare and profit are maximized by complete separation. That is, every type is in a group of his own, so agents associate with those just like themselves and ignore everyone else. Conversely, when a quality function exhibits positive returns to scale, in that merging two groups raises the average quality, then matching is assortative (i.e. groups are connected) and there tends to be some pooling.

The paper also examines the principal's problem when group formation is costly. This setting introduces a new factor, the *appropriability effect*, according to which a welfare-maximizing principal may invest more in group formation than a profit-maximizing principal. Intuitively, a profit-maximizer cannot appropriate agents' consumer surplus and may opt for larger groups than is optimal. Nevertheless, under positive returns to scale and the usual monotone hazard rate condition, the segregation effect dominates the appropriability effect and groups are smaller under profit-maximization.

We also investigate how welfare- and profit-maximizing group structures change as agents' types increase. This is motivated by Lazear (2001) who argues that more able

students tend to be in larger classes. In our model, when group formation is costless, we also find that an increase in all agents' types lead groups to become larger, albeit for a very different reason. The intuition behind our result is that the ratio between the highest and lowest types in a group declines as everyone's type rises. This means a group split, which helps the high types but hurts the low types, becomes less desirable. In comparison, Lazear's finding derives from the specifics of the multiplicative quality model, under which returns to scale increase in agents' types.

1.2 *Theoretical literature*

It is helpful to break the peer group literature into three branches.

The first branch considers a single principal with perfect information about agents' characteristics. In their classic paper, Arnott and Rowse (1987) analyze the socially optimal way to break students into N groups in the presence of peer effects. A student's utility is a function of his ability, the mean ability of the other students in the class, and educational expenditure. Using a Cobb–Douglas quality function, the authors obtain sufficient conditions for assortative matching and computationally solve several examples. Lazear (2001) considers a highly tractable model where each student is disruptive with probability p. If there are m students in the class who act independently of each other then the class is attentive proportion $(1-p)^m$ of the time. Lazear shows that a welfare-maximizing school increases class size as p increases and, in a two-type model, segregates students by type.

In the second branch, there is a single principal with imperfect information about agents' characteristics. Helsley and Strange (2000) analyze common interest developments with social interactions. Agents, who vary in their type, choose whether to stay in the public sector or join a single private community, and subsequently choose an action. An agent's utility then depends upon their action, their type, and the mean action of those in their community. Helsley and Strange allow the private community to choose both a minimum required action and an entry price. In a numerical example they show fewer people secede from the public sector when the community is profit-maximizing, in a similar spirit to our exclusion effect.

The two closest papers to the current one both consider a principal who price discriminates between agents by sorting them into different groups of different qualities. Rayo (2003) considers a one-sided matching problem, similar to ours, where the principal breaks the agents up into groups. Rayo uses the average-quality function and investigates the role of non-monotone marginal revenue functions (see Section 5.4). Damiano and Li (2007) analyze a two-sided matching market where the principal can discriminate between different sides of the market and between different groups. They derive necessary and sufficient conditions on the distribution of types for full separation. In comparison to these papers, we allow for a more flexible form of peer effects that encompasses a number of different models as special cases.[2]

The third branch analyzes competition between peer groups. Epple and Romano

[2] In a methodologically separate line of work, Moldovanu et al. (2007) considers the optimal grouping of a finite number of agents who care about their status.

(1998) consider a model of private school competition, where agents differ in their income and ability, both of which are publicly observable. Epple and Romano show that monopolistic competition between schools with fixed costs leads to stratification of the market where poor talented agents attend the same schools as wealthy untalented agents. Caucutt (2002) introduces educational expenditure and shows that complete segregation may not be desirable, even without fixed costs of setting up schools. Intuitively, a school can keep its quality constant by lowering its expenditure on teachers but recruiting a few talented students. Farrell and Scotchmer (1988) analyze a perfect-information model where agents form groups and then split the output of each group between the members. In the unique core allocation, groups are connected and are too finely segregated: intuitively, high types do not internalize their positive externality on low types.[3] Finally, in a model with imperfectly observed types, Damiano and Li (2008) analyze the competition between two matchmakers who each simultaneously choose a single group-entry price.

1.3 *Empirical literature*

A large empirical literature seeks to estimate peer technology. In the classroom, Henderson et al. (1978) study how a student's exam performance is affected by the ability of their peers. The paper has three main findings. First, the magnitude of these peer effects is substantial. Moving a student from a weak class to a strong class can increase their overall rank from the 50$^{\text{th}}$ percentile to the 20$^{\text{th}}$ percentile. Second, test scores are a concave function of mean class ability. This result is consistent with a generalized average quality function where $\phi(\cdot)$ is concave (see Section 5.4), implying positive returns to scale and suggesting that some mixing of abilities is optimal (see Propositions 2–3). Third, there are insignificant interaction effects between a student's performance and the mean ability of the class. This implies that, if all students care equally about their test scores, then it is impossible to separate different types of agents (see Lemma 2). However, if high ability students care more about their test scores than low ability students, then separation can be sustained. Subsequent classroom studies have added to these results. Hanushek et al. (2003) find that the variance of ability within a classroom has no clear effect on a student's performance, suggesting the mean is a reasonable summary statistic, as in the generalized average quality model. Hanushek (1999) considers the effect of class size, reporting that, while the desirability of small classes may seem obvious, the evidence seems to find beneficial effects only in certain environments

In a study of college roommates, Zimmerman (2003) examines the impact of one roommate's performance on the other. He finds that bad students have a bigger effect on their roommates than good students, suggesting negative returns to scale, and implying that bad students should be segregated (see Proposition 2). However, in a similar study, Sacerdote (2001) finds that good students have a bigger effect on their roommates than

[3] One can view these papers as applications of club theory (e.g. Scotchmer 2002). For example, Ellickson et al. (1999) show that, when a full set of group- and type-dependent prices is available, then the competitive equilibria are Pareto efficient and coincide with the core. On a more applied level, Nechyba (2000) and Benabou (1993) look at competition across schools and cities.

bad students, suggesting positive returns to scale, and implying that some mixing of abilities is optimal (see Propositions 2–3).

Mas and Moretti (forthcoming) find substantial peer effects in the workplace. They show that a 10% increase in the productivity of a worker's colleagues leads that worker's productivity to rise by 1.5%. They also find good peers have a stronger effect on poor workers. Since there are only small numbers of workers, it is unclear whether this is due to positive returns to scale in the production process (good workers have more effect on the environment than poor workers) or due to interaction effects (poor workers gain more from an improvement in the environment).

Looking across these studies, it seems that the peer technology can vary greatly with the environment. This observation has two important implications. First, it is important to derive results, like the segregation effect, that do not depend on the exact nature of the peer effects. Second, theory should identify which aspects of the peer technology are critical for a given result, rather than working with a single functional form, which contains many hidden assumptions. This approach both helps us categorize different classes of peer technologies, and helps us understand what to look for in the data.

2. Two-type example

There are equal numbers of two types of agents, $\theta_H > \theta_L$, where an agent's type describes his willingness to pay for quality. The utility of type θ_i who is assigned to a group of quality $\mathcal{Q}(\theta_i)$ and pays price $y(\theta_i)$ is given by $u(\theta_i) = \theta_i \mathcal{Q}(\theta_i) - y(\theta_i)$, for $i \in \{L, H\}$. The quality of a group is determined by the types of its members. A group consisting only of θ_H agents has quality Q_H; a group consisting only of θ_L agents has quality Q_L; and a group consisting of both types has quality Q_{LH}.[4] An agent's outside option is 0. Finally, we suppose that agents are small, so no individual agent can affect the quality of a group.

The principal posts anonymous group-entry prices and lets agents self-select into the different groups. This means that, in order to stop the high type copying the low type, we must have $\mathcal{Q}(\theta_H) \geq \mathcal{Q}(\theta_L)$ (the monotonicity condition). Consequently, the principal can separate the agents if and only if $Q_H \geq Q_L$; otherwise a high type would enter the low type's group rather than his own.[5]

2.1 *Segregation effect (see Section 5)*

Let us first consider the principal's incentive to separate the two types of agents. For simplicity, assume that $2\theta_L \geq \theta_H$ and that the principal does not exclude either type. Utility is quasi-linear, so welfare equals $\theta_L \mathcal{Q}(\theta_L) + \theta_H \mathcal{Q}(\theta_H)$. A welfare-maximizing principal would therefore like to separate the agents when

$$\theta_H Q_H + \theta_L Q_L \geq \theta_H Q_{LH} + \theta_L Q_{LH}. \tag{1}$$

[4] We allow the principal to choose between pooling and separation, but do not allow for stochastic mechanisms. While this is restrictive in a two-type model, it is less problematic in a many-type model. The results are also robust to stochastic mechanisms: see footnote 10.

[5] The principal is female, while agents are male.

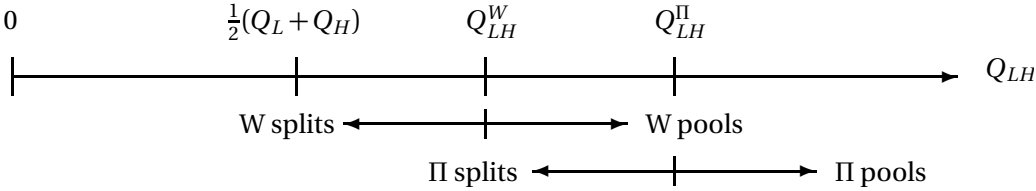

FIGURE 1. Two-type model with $Q_H \geq Q_L$. This figure shows (a) there is more separation under profit-maximization than welfare-maximization; and (b) under either objective, separation is optimal if the quality function satisfies negative returns to scale.

Define Q^W_{LH} as the pooling quality that equates the two sides of (1). If $Q_H < Q_L$, the principal can only pool the agents. If $Q_H \geq Q_L$, then the principal separates the agents when $Q_{LH} \leq Q^W_{LH}$. We say the quality function exhibits *negative returns to scale* if $Q_{LH} \leq \frac{1}{2}(Q_H + Q_L)$, so that pooling the agents lowers the average quality. Since $Q^W_{LH} \geq \frac{1}{2}(Q_H + Q_L)$, it follows that separation is optimal if the quality function is increasing, $Q_H \geq Q_L$, and satisfies negative returns to scale.

A profit-maximizing principal maximizes total payments, $y(\theta_L)+y(\theta_H)$. If the principal pools both types, she charges $y(\theta_L) = y(\theta_H) = \theta_L Q_{LH}$ in order to fully extract from the low type, θ_L. On the other hand, if the principal separates both types, she charges $y(\theta_L) = \theta_L Q_L$ to the low group and $y(\theta_H) = \theta_L Q_H - (\theta_H - \theta_L)Q_L$ to the high group. Under these prices, the low type is just willing to join the low group, while the high type is indifferent between joining the high and low groups. Putting this together, the profit-maximizing principal would like to separate the agents when

$$\theta_H Q_H + (2\theta_L - \theta_H)Q_L \geq 2\theta_L Q_{LH}. \tag{2}$$

Define Q^Π_{LH} as the pooling quality that equates the two sides of (2). If $Q_H < Q_L$, then the principal can only pool the agents. If $Q_H \geq Q_L$, then $Q^\Pi_{LH} \geq Q^W_{LH}$, so a profit-maximizing principal is more willing to separate the agents than a welfare-maximizing principal (see Figure 1). Intuitively, by separating high and low types, the good agents become very keen to avoid the bad agents and can be forced to pay higher prices. Notice that this *segregation effect* requires no assumptions about the structure of qualities (Q_L, Q_H, Q_{LH}): the fact that $\mathcal{Q}(\theta_H) \geq \mathcal{Q}(\theta_L)$ follows from the endogenous self-selection constraint.

2.2 *Exclusion effect (see Section 6)*

If $2\theta_L < \theta_H$, then the profit-maximizing principal may wish to exclude the low types in order to increase revenue. To see this, consider the case where $Q_H \geq Q_L$. The welfare-maximizing principal never excludes any type of agent, and separates the two types if (1) holds. In contrast, the profit-maximizing principal may wish to exclude the low type, enabling her to charge $y(\theta_H) = \theta_H Q_H$ to the remaining high types. She therefore wishes to separate the two types if

$$\theta_H Q_H + \max\{2\theta_L - \theta_H, 0\}Q_L \geq 2\theta_L Q_{LH}. \tag{3}$$

As above, (1) implies (3). This shows that the segregation effect extends to the case where we allow exclusion. Moreover, a profit-maximizing principal is more willing to exclude agents than a welfare-maximizing principal. This *exclusion effect* is analogous to the standard monopoly distortion: by cutting out low types the principal increases the price she can charge the high types.

2.3 *Appropriability effect (see Section 7)*

So far we have assumed that splitting the agents into two groups is free of charge. Costly group formation introduces a third effect. To illustrate, let us assume that $Q_H \geq Q_L$. Using (1), the benefit of separation for a welfare-maximizing principal is

$$\theta_H(Q_H - Q_{LH}) + \theta_L(Q_L - Q_{LH}).$$

If $2\theta_L \geq \theta_H$, (2) implies that the benefit of separation for a profit-maximizing principal is

$$\theta_H(Q_H - Q_{LH}) + (2\theta_L - \theta_H)(Q_L - Q_{LH}).$$

Hence, if group formation is costly and there are very strong negative returns to scale, in that $Q_L \geq Q_{LH}$, then the welfare-maximizing principal is willing to pay more to separate the agents than the profit-maximizing principal. This *appropriability effect* is caused by the profit-maximizing principal's inability to appropriate the agents' consumer surplus. However, when there are positive returns to scale, in that $Q_{LH} \geq \frac{1}{2}(Q_H + Q_L)$, then the segregation effect outweighs the appropriability effect and the profit-maximizing principal is more likely to separate the groups.

3. Basic model

Agents' preferences A single principal faces a continuum of agents with privately known willingness to pay $\theta \in [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}_+$, where we allow $\overline{\theta} = \infty$. Types are distributed according to the positive density $f(\theta)$ with distribution function $F(\theta)$. If agent θ is assigned to a group of quality Q and pays price y, he obtains utility

$$u = \theta Q - y.$$

If an agent is assigned to no group, he obtains zero utility.

Mechanism We are interested in a model where the principal chooses prices and agents self-select into groups. Appealing to the revelation principle,[6] we analyze the direct revelation mechanism $\langle \mathcal{G}, y \rangle$ consisting of a group structure $\mathcal{G}$, defined below, and payments $y : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$.

[6]The revelation principle says that, given any equilibrium in the price-setting game, there exists a corresponding direct revelation mechanism such that all agents accept the mechanism (individual rationality) and announce their types truthfully (incentive compatibility).

Groups The principal breaks the agents into groups. A *group* $G \subset [\underline{\theta}, \overline{\theta}]$ is a Borel set. A *group structure* $\mathcal{G}$ is a collection of nonintersecting groups whose union equals $[\underline{\theta}, \overline{\theta}]$. Given a mechanism $\langle \mathcal{G}, y \rangle$, an agent who declares $\hat{\theta}$ is assigned to the corresponding group in $\mathcal{G}$. Three definitions are useful. Taking two group structures, $\mathcal{G}_C$ and $\mathcal{G}_F$, we write $\mathcal{G}_F \succcurlyeq \mathcal{G}_C$ if $\mathcal{G}_F$ is finer than $\mathcal{G}_C$. Two groups, G and G', *overlap* if there exists $\theta_H > \theta_M > \theta_L$, such that $\theta_H, \theta_L \in G$ and $\theta_M \in G'$. Finally, G_H *larger* than G_L if $\theta_H \geq \theta_L$ for all $\theta_H \in G_H$ and $\theta_L \in G_L$

Peer technology Each group $G \subset [\underline{\theta}, \overline{\theta}]$ is associated with a quality $Q(G) > 0$. Given a group structure $\mathcal{G}$, let $\mathcal{Q}(\theta, \mathcal{G})$ denote the quality of type θ's group, and assume $\mathcal{Q}$ is integrable on $[\underline{\theta}, \overline{\theta}]$.[7] A quality function $Q(G)$ is *weakly increasing* in G if $Q(G_H) \geq Q(G_L)$ whenever G_H is larger than G_L. Similarly, $Q(G)$ is *weakly decreasing* in G if $Q(G_H) \leq Q(G_L)$ whenever G_H is larger than G_L.

Agents' problem Given a mechanism $\langle \mathcal{G}, y \rangle$, an agent of type θ declares that he is type $\hat{\theta}$, receives quality $\mathcal{Q}(\hat{\theta}, \mathcal{G})$, and makes payment $y(\hat{\theta})$. Since there is a continuum of agents in every group, the quality of an agent's group depends on his declaration, but not his type. His utility is then

$$u(\theta, \hat{\theta}) = \theta \mathcal{Q}(\hat{\theta}, \mathcal{G}) - y(\hat{\theta}).$$

Principal's problem The principal chooses a mechanism $\langle \mathcal{G}, y \rangle$ to maximize welfare/profits subject to individual rationality (each agent receives positive utility) and incentive compatibility (each agent declares his type truthfully).

Other definitions We say a function $\phi : \mathbb{R} \to \mathbb{R}$ is *quasi-increasing* if $\phi(x_L) \geq 0$ implies $\phi(x_H) \geq 0$ for $x_H > x_L$.

Some remarks are in order. First, we allow for a large range of quality functions, $Q(G)$, subsuming those used in a number of previous papers (see Section 5.4). This level of generality is important since the peer technology depends on the specific environment and is hard to quantify in any given application.

Second, we do not insist that groups be connected. This is important because the optimal group structure may place agents with a wide range of abilities in the same group, as suggested by the empirical work of Henderson et al. (1978) and Mas and Moretti (forthcoming). For a theoretical illustration, see Example G in Section 5.4.

Third, we assume that the principal places each agent into a group. That is, we suppose it is not optimal for the principal to exclude any types of agents. This assumption is for simplicity: we extend the analysis in Section 6.

4. Agents' problem

Define equilibrium utility to be $U(\theta) = u(\theta, \theta)$.

[7] For example, $\mathcal{Q}(\theta, \mathcal{G})$ is integrable if $Q(G)$ is bounded.

LEMMA 1. *A mechanism $\langle \mathcal{G}, y \rangle$ is incentive compatible and individually rational if and only if*

(i) Utility is given by

$$U(\theta) = \int_{\underline{\theta}}^{\theta} \mathcal{Q}(s, \mathcal{G})\, ds + U(\underline{\theta}). \tag{4}$$

(ii) The lowest type obtains $U(\underline{\theta}) \geq 0$.

(iii) The monotonicity condition holds. That is, $\mathcal{Q}(\theta, \mathcal{G})$ is increasing in θ.

PROOF. Since $\mathcal{Q}(\theta, \mathcal{G})$ is integrable, Milgrom and Segal (2002, Corollary 1) shows that incentive compatibility implies (4). The rest of the proof is the same as Mas-Colell et al. (1995, Proposition 23.D.2). □

LEMMA 2. *In any incentive compatible group structure,*

(i) any overlapping groups have the same quality

(ii) if $Q(G)$ is weakly decreasing then every agent is in a group of the same quality.

This result follows from the monotonicity condition (Lemma 1(iii)).

Lemma 2(i) says that while groups do not have to be connected, any overlapping groups must have the same quality. Lemma 2(ii) says that the principal cannot separate different types when the agents who generate high quality have a low willingness to pay. Separation may therefore be difficult with some conspicuous goods, where agents seek to signal a certain image. For example, the consumers who generate Harley-Davidson's reputation are unlikely to have the highest incomes. Similarly, the supporters with the highest willingness to pay for football tickets may not create the best atmosphere.

5. THE SEGREGATION EFFECT

5.1 *Principal's problem*

Welfare equals the sum of utilities plus transfers,

$$W = E[\theta\, \mathcal{Q}(\theta, \mathcal{G})]. \tag{5}$$

Integrating utility (4) by parts, consumer surplus is

$$E[U(\theta)] = E\left[\frac{1 - F(\theta)}{f(\theta)}\, \mathcal{Q}(\theta, \mathcal{G})\right] + U(\underline{\theta}). \tag{6}$$

Profit equals welfare (5) minus consumer surplus (6). The profit-maximizing principal sets prices so that the lowest type's individual rationality constraint binds, $U(\underline{\theta}) = 0$. Profit is then given by

$$\Pi = E\left[MR(\theta)\, \mathcal{Q}(\theta, \mathcal{G})\right],$$

where marginal revenue is defined by

$$MR(\theta) := \theta - \frac{1-F(\theta)}{f(\theta)}.$$

Welfare and profit can be thus combined into a single objective:

$$H = E\left[h(\theta)\mathcal{Q}(\theta,\mathcal{G})\right] \tag{7}$$

where $h(\theta) \in \{\theta, MR(\theta)\}$. Let Γ be the set of group structures that satisfy the monotonicity condition (Lemma 1(iii)). The principal's problem is then to choose $\mathcal{G} \in \Gamma$ to maximize (7). Assume a solution to this problem exists.[8]

5.2 *Welfare- and profit-maximization*

An *interval partition* is a collection of intervals (connected sets) whose union equals $[\underline{\theta}, \overline{\theta}]$. For a fixed group structure $\mathcal{G}$, let $\mathcal{I}(\mathcal{G})$ be the finest connected coarsening of $\mathcal{G}$, formed by merging all overlapping groups. Formally, $\mathcal{I}(\mathcal{G})$ is the join (coarsest common refinement) of all interval partitions formed by taking unions of groups in $\mathcal{G}$. Lemma 2(i) then implies that quality is constant over each $I \in \mathcal{I}(\mathcal{G})$.

LEMMA 3. *$\mathcal{G}_F \succcurlyeq \mathcal{G}_C$ implies $\mathcal{I}(\mathcal{G}_F) \succcurlyeq \mathcal{I}(\mathcal{G}_C)$.*

PROOF. Let Λ_C (resp. Λ_F) be the set of interval partitions formed by taking unions of groups in $\mathcal{G}_C$ (resp. $\mathcal{G}_F$). Pick $\mathcal{I} \in \Lambda_C$. Since $\mathcal{G}_F \succcurlyeq \mathcal{G}_C$, $\mathcal{I}$ can also be formed by taking unions of groups in $\mathcal{G}_F$. That is, $\Lambda_C \subset \Lambda_F$. As a consequence,

$$\mathcal{I}(\mathcal{G}_F) = \vee_{\mathcal{I}\in\Lambda_F}\mathcal{I} \succcurlyeq \vee_{\mathcal{I}\in\Lambda_C}\mathcal{I} = \mathcal{I}(\mathcal{G}_C)$$

where $\vee$ denotes the join. □

ASSUMPTION (MON). $[1-F(\theta)]/\theta f(\theta)$ is decreasing in θ.

This assumption implies that $MR(\theta)$ is quasi-increasing. It is weaker than the usual hazard rate condition (see Section 7).

Proposition 1 is our main result. It states that, starting from a welfare-maximizing group structure, profit is reduced by merging groups further. That is, the profit-maximizing group structure is no coarser than the welfare-maximizing group structure.

PROPOSITION 1 (Segregation effect). *Suppose (MON) holds and $MR(\underline{\theta}) \geq 0$. Pick any welfare-maximizing solution, $\mathcal{G}^W$. Then $\Pi(\mathcal{G}^W) \geq \Pi(\mathcal{G})$ for all $\mathcal{G} \in \Gamma$ that are coarser than $\mathcal{G}^W$.*

[8] A solution is guaranteed to exist for a discrete version of the problem, where we break $[\underline{\theta}, \overline{\theta}]$ into a finite interval partition $\mathcal{P}$, and a group is restricted to be measurable with respect to $\mathcal{P}$. While this discretization rules out the use of calculus (e.g. Proposition 3), it does not affect our other results.

Proof. Suppose $\mathcal{G}^W$ maximizes welfare and fix $\mathcal{G} \in \Gamma$ such that $\mathcal{G} \preccurlyeq \mathcal{G}^W$. Since $\mathcal{G}^W$ is welfare-maximizing, $E[\theta \Delta\mathcal{Q}(\theta)] \geq 0$, where $\Delta\mathcal{Q}(\theta) := \mathcal{Q}(\theta, \mathcal{G}^W) - \mathcal{Q}(\theta, \mathcal{G})$. Define $\mathcal{I}^*$ to be the coarsest interval partition on which $\Delta\mathcal{Q}(\theta)$ is quasi-increasing (see Figure 2). Applying Lemma 3, $\mathcal{I}(\mathcal{G}) \preccurlyeq \mathcal{I}(\mathcal{G}^W)$. Since $\mathcal{Q}(\theta, \mathcal{G})$ is constant on each $I \in \mathcal{I}(\mathcal{G})$, monotonicity implies that $\Delta\mathcal{Q}(\theta)$ is increasing on each $I \in \mathcal{I}(\mathcal{G})$, implying $\mathcal{I}^* \preccurlyeq \mathcal{I}(\mathcal{G})$. Moreover, each $I \in \mathcal{I}^*$ has positive measure since $\Delta\mathcal{Q}(\theta) = 0$ on single points in $\mathcal{I}(\mathcal{G})$. The proof is now based on two steps.

For the first step, we claim that $E[\theta \Delta\mathcal{Q}(\theta) \,|\, \mathcal{I}^*] \geq 0$.[9] To see this suppose, by contradiction, that $E[\theta \Delta\mathcal{Q}(\theta) \,|\, \mathcal{I}^*] < 0$ on some set $I \in \mathcal{I}^*$. Then define a new group structure, $\mathcal{G}'$, equal to $\mathcal{G}$ on I and $\mathcal{G}^W$ elsewhere. This new structure has two properties. First, $\mathcal{G}'$ has higher welfare than $\mathcal{G}^W$, $E[\theta\, \mathcal{Q}(\theta, \mathcal{G}')] > E[\theta\, \mathcal{Q}(\theta, \mathcal{G}^W)]$. Second, $\mathcal{G}' \in \Gamma$, which we verify below. Together, these contradict the welfare-optimality of $\mathcal{G}^W$.

Let us now verify that $\mathcal{G}' \in \Gamma$. The interval partition $\mathcal{I}^*$ has the key property that $\Delta\mathcal{Q}(\theta)$ goes from negative to positive on each $I^* \in \mathcal{I}^*$. Formally, there exists $\theta_1(I^*) \in I^*$ such that $\Delta\mathcal{Q}(\theta) \leq 0$ on $\{\theta \in I^* : \theta \leq \theta_1(I^*)\}$, except possibly for the lowest interval. Similarly, there exists $\theta_2(I^*) \in I^*$ such that $\Delta\mathcal{Q}(\theta) \geq 0$ on $\{\theta \in I^* : \theta \geq \theta_2(I^*)\}$, except possibly for the highest interval. To show $\mathcal{Q}(\theta, \mathcal{G}')$ is increasing, pick $\theta_H > \theta_L$ and denote the respective interval partitions $I_H, I_L \in \mathcal{I}^*$. If $I_H = I_L$, then $\mathcal{Q}(\theta_H, \mathcal{G}') \geq \mathcal{Q}(\theta_L, \mathcal{G}')$ follows from the monotonicity of $\mathcal{Q}(\theta, \mathcal{G})$ and $\mathcal{Q}(\theta, \mathcal{G}^W)$. Next, suppose $I_H \neq I_L$ and fix $\theta' \in \{\theta \in I_H : \theta \leq \theta_H, \theta \leq \theta_1(I_H)\}$ and $\theta'' \in \{\theta \in I_L : \theta \geq \theta_L, \theta \geq \theta_2(I_L)\}$. Then,

$$\mathcal{Q}(\theta_H, \mathcal{G}') \geq Q(\theta', \mathcal{G}') \geq \mathcal{Q}(\theta', \mathcal{G}^W) \geq \mathcal{Q}(\theta'', \mathcal{G}^W) \geq \mathcal{Q}(\theta'', \mathcal{G}') \geq \mathcal{Q}(\theta_L, \mathcal{G}').$$

The first, third, and fifth inequalities come from monotonicity. The second and fourth inequalities come from the above properties of $\mathcal{I}^*$. Hence $\mathcal{G}' \in \Gamma$, as required.

For the second step, index the objective function $h(\theta, t)$ so that $h(\theta, 1) = MR(\theta) \geq 0$ and $h(\theta, 0) = \theta$. Under (MON), the function $h(\theta, t) \geq 0$ is log-supermodular. Since $\Delta\mathcal{Q}(\theta)$ is quasi-increasing on each $I^* \in \mathcal{I}^*$, Lemma 5(i) in Section A.1 states that $E[h(\theta, t)\Delta\mathcal{Q}(\theta) \mid \mathcal{I}^*]$ is quasi-increasing in t. Thus $E[\theta \Delta\mathcal{Q}(\theta) \mid \mathcal{I}^*] \geq 0$ implies that $E[MR(\theta)\Delta\mathcal{Q}(\theta) \mid \mathcal{I}^*] \geq 0$. Integrating over θ, we have $E[MR(\theta)\Delta\mathcal{Q}(\theta)] \geq 0$. That is, $\Pi(\mathcal{G}^W) \geq \Pi(\mathcal{G})$. □

Corollary 1. *Suppose the welfare- and profit-maximizing group structures satisfy $\mathcal{G}^W \preccurlyeq \mathcal{G}^\Pi$. Then $E[\phi \circ \mathcal{Q}(\theta, \mathcal{G}^W)] \geq E[\phi \circ \mathcal{Q}(\theta, \mathcal{G}^\Pi)]$ for any increasing concave function $\phi : \mathbb{R} \to \mathbb{R}$.*

The proof of this corollary is given in Section A.2.

Proposition 1 says that the profit-maximizing group structure is no coarser than the welfare-maximizing group structure. If these optimal group structures are ordered then Corollary 1 says that profit-maximization induces a distribution of quality levels that has a lower mean and tends to be more dispersed. In the school example, if one interprets $\mathcal{Q}(\theta, \mathcal{G})$ as the exam scores of agent θ, then Corollary 1 yields testable implications of the theory.

[9] Notation: the function $E[\theta \Delta\mathcal{Q}(\theta) \,|\, \mathcal{I}^*] : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$ maps each type into its conditional expectation.

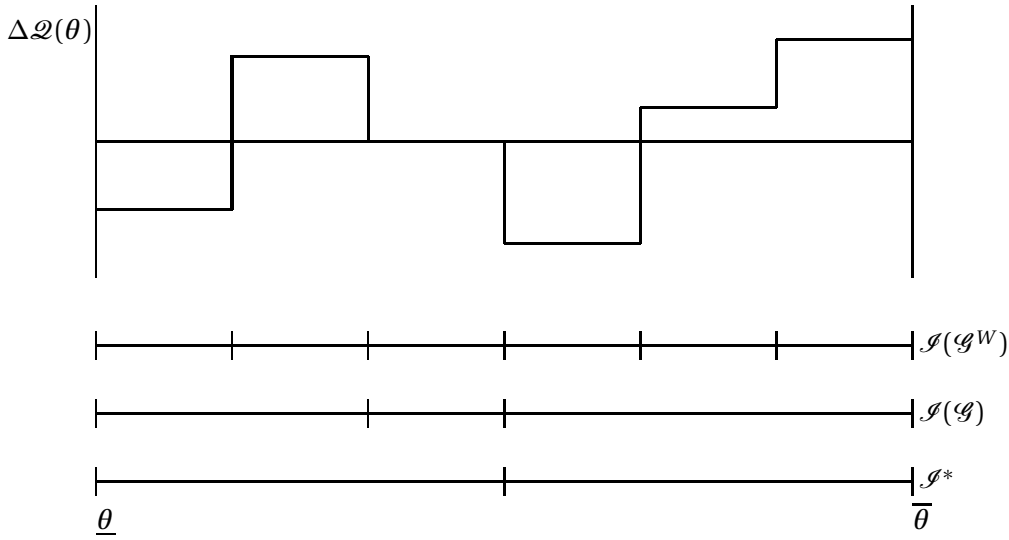

FIGURE 2. Illustration of sets in proof of Proposition 1.

The idea behind the segregation effect is that, under (MON), $MR(\theta_H)/MR(\theta_L) \geq \theta_H/\theta_L$ for $\theta_H > \theta_L$, so a profit-maximizing firm puts relatively more weight on the preferences of high types than does the social planner. This means a profit-maximizing firm is more likely to split up a group, which helps the high types and hurts the low types. Intuitively, by introducing extra segregation the principal raises the cost of pretending to be a lower type and reduces consumer surplus. That is, by separating good and bad agents, the good agents become very keen to avoid the bad groups and can be forced to pay higher prices.

As stated in the Introduction, we make no assumption about the nature of peer effects. This is important because peer technology differs greatly across environments. Instead, the result uses only the monotonicity condition that comes endogenously from the agents' self-selection constraints.

Proposition 1 states that, starting from the welfare-maximizing outcome $\mathcal{G}^W$, profit is not increased by merging groups. One can also show that, starting from the profit-maximizing outcome $\mathcal{G}^\Pi$, welfare is not increased by splitting groups. The proof is the same, although one should use Lemma 5(ii) rather than Lemma 5(i).[10]

Proposition 1 does have one limitation in that the welfare- and profit-maximizing groups may not be ordered. Intuitively, this comes from the fact that different ways of dividing a group are likely to be substitutes, rather than complements. One should therefore view the result as saying that, if we start from the welfare-maximizing group

[10]One can also extend the model to allow for multidimensional types and stochastic mechanisms. To do this, endow each agent with a two-dimensional type (θ, q), where θ is the willingness to pay and q determines quality, and define a group to be a measure over (θ, q). In this more general environment, the segregation effect (Proposition 1) continues to apply. Intuitively, an agent's utility depends on his payoff type θ, but not his quality type q. Hence the the principal can screen only on θ, and the two-dimensional screening problem collapses to a one-dimensional screening problem (Board 2007, Section 9.2).

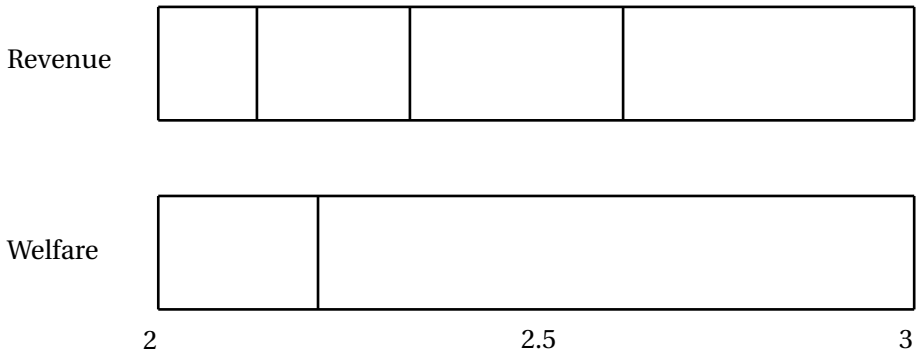

FIGURE 3. Optimal group formation: maxmin quality. This figure shows the profit and welfare-maximizing group structures where $Q(G) = 0.55\sup(G) + 0.45\inf(G)$ and $\theta \sim U[2,3]$. See Section 5.4 and Example 5 for more details.

structure, then separating groups may increase profit, but merging groups does not. Figure 3 shows that the spirit of the result may remain true even if the optimal solutions are not ordered.

The following example shows that the (MON) condition is tight.

EXAMPLE 1 (Pareto distribution). Suppose $\theta \sim \text{Par}(\alpha,\beta)$, so that $f(\theta) = \alpha\beta^{\alpha}\theta^{-(\alpha+1)}$. In this case, (MON) holds with equality and profit is $(1-\alpha^{-1})E[\theta\mathcal{Q}(\theta,\mathcal{G})]$. Consequently, the welfare- and profit-maximizing group choices coincide. ◊

5.3 *Group structure and returns to scale*

This section analyzes how peer technology affects the optimal group structure. Define the *returns to scale* by

$$R(\mathcal{G}_C,\mathcal{G}_F) := E[\mathcal{Q}(\theta,\mathcal{G}_C)] - E[\mathcal{Q}(\theta,\mathcal{G}_F)]$$

for $\mathcal{G}_F \succcurlyeq \mathcal{G}_C$.

DEFINITION 1. The quality function exhibits

(i) *positive returns to scale* (PRS) if $R(\mathcal{G}_C,\mathcal{G}_F) \geq 0$ for all $\mathcal{G}_F \succcurlyeq \mathcal{G}_C$

(ii) *negative returns to scale* (NRS) if $R(\mathcal{G}_C,\mathcal{G}_F) \leq 0$ for all $\mathcal{G}_F \succcurlyeq \mathcal{G}_C$.

Under NRS, merging groups lowers the average quality. Under PRS, merging groups raises the average quality. Which case is appropriate depends upon the application and the interpretation of a group. To illustrate, consider the school example. If one interprets a group as a class, then merging two classes into one large class is likely to harm all students' performances. This suggests that the quality function satisfies NRS. On the other hand, if one fixes the class size and interprets a group as an entire school, then the good students may help the poor students more than the poor students harm the good students (e.g. Henderson et al. 1978). In this case, the quality function satisfies PRS.

Recall $h(\theta) \in \{\theta, MR(\theta)\}$. The following result assumes $h(\theta)$ is increasing. It therefore applies to the welfare-maximization problem and, when $MR(\theta)$ is increasing, to the profit-maximization problem.

PROPOSITION 2. *Assume $h(\theta)$ is positive and increasing, and $Q(G)$ is weakly increasing in G.*

(i) Under NRS, the optimum is attained by full separation, where each type is in a group of their own.

(ii) Under PRS, the optimum is attained by an interval partition, where all groups are connected.

PROOF. (i) We prove a more general result: Suppose $h(\theta)$ is positive and increasing, and that NRS holds. Then, for any $\mathscr{G}_C, \mathscr{G}_F \in \Gamma$ such that $\mathscr{G}_F \succcurlyeq \mathscr{G}_C$, the principal prefers $\mathscr{G}_F$ to $\mathscr{G}_C$. If $Q(G)$ is weakly increasing then full separation satisfies monotonicity and therefore maximizes the principal's payoff.

Pick $\mathscr{G}_C, \mathscr{G}_F \in \Gamma$ such that $\mathscr{G}_F \succcurlyeq \mathscr{G}_C$, and denote $\Delta\mathscr{Q}(\theta) := \mathscr{Q}(\theta, \mathscr{G}_F) - \mathscr{Q}(\theta, \mathscr{G}_C)$. Let $\mathscr{I}^*$ be the coarsest interval partition on which $\Delta\mathscr{Q}(\theta)$ is quasi-increasing. By Lemma 3, $\mathscr{I}(\mathscr{G}_F) \succcurlyeq \mathscr{I}(\mathscr{G}_C) \succcurlyeq \mathscr{I}^*$. We claim that NRS implies

$$E[\Delta\mathscr{Q}(\theta) \mid \mathscr{I}^*] \geq 0. \tag{8}$$

To see this pick $I^* \in \mathscr{I}^*$ and let $\mathscr{G}'$ equal $\mathscr{G}_F$ on I^* and equal $\mathscr{G}_C$ elsewhere. First, $\mathscr{G}' \in \Gamma$, as in the proof of Proposition 1. Second, since $\mathscr{G}' \succcurlyeq \mathscr{G}_C$, NRS implies

$$E[\Delta\mathscr{Q}(\theta) \mid I^*] = E[\mathscr{Q}(\theta, \mathscr{G}')] - E[\mathscr{Q}(\theta, \mathscr{G}_C)] \geq 0$$

as required. This result implies that

$$E[h(\theta)\Delta\mathscr{Q}(\theta) \mid \mathscr{I}^*] \geq E[h(\theta) \mid \mathscr{I}^*]E[\Delta\mathscr{Q}(\theta) \mid \mathscr{I}^*] \geq 0,$$

where the first inequality comes from the fact that an increasing function and a quasi-increasing function have positive covariance (e.g. Persico 2000, Lemma 1), and the second from (8). Integrating over θ, $\mathscr{G}_F$ yields a higher payoff than does $\mathscr{G}_C$.

(ii) Suppose there is PRS. Pick an arbitrary group structure, $\mathscr{G} \in \Gamma$. Let $\mathscr{I}(\mathscr{G})$ be the finest connected coarsening of $\mathscr{G}$. Since $Q(G)$ is weakly increasing, we have $\mathscr{I}(\mathscr{G}) \in \Gamma$. Moreover, PRS implies that merging increases group quality, so $E[h(\theta)\mathscr{Q}(\theta, \mathscr{I}(\mathscr{G}))] \geq E[h(\theta)\mathscr{Q}(\theta, \mathscr{G})]$. □

Proposition 2 shows that if the quality function exhibits negative returns to scale, then there is full separation. Proposition 3 is a partial converse: it shows that if the quality function exhibits positive returns to scale that are of first-order, then there is some pooling. Define the *local returns to scale* at θ by

$$R_\theta(\epsilon) := Q(G_\epsilon) - E[Q(\{x\}) \mid G_\epsilon] \tag{9}$$

where $G_\epsilon := [\theta, \theta + \epsilon]$ and $Q(\{x\})$ is the quality of x's group under full separation. We say that a group structure exhibits pooling at θ if we cannot find an open set $(\theta - \epsilon, \theta + \epsilon)$ over which there is full separation.

PROPOSITION 3. *Assume $h(\theta)$ is strictly positive and $Q(G)$ is weakly increasing and exhibits PRS. In addition, suppose that $h(\theta)$ and $Q(\{\theta\})$ are continuously differentiable. Then any optimal group structure exhibits pooling at $\theta \in (\underline{\theta}, \overline{\theta})$ if there exist $\overline{\epsilon}, \delta > 0$ such that*

$$h(\theta)R_\theta(\epsilon) \geq \left[\tfrac{1}{12}h'(\theta)Q'(\{\theta\}) + \delta\right]\epsilon^2 \qquad \text{for } \epsilon \in (0, \overline{\epsilon}]. \tag{10}$$

This is satisfied if $R'_\theta(0) := \lim_{\epsilon \to 0} R_\theta(\epsilon)/\epsilon > 0$.

PROOF. Since $h(\theta)$ is positive and $Q(G)$ is weakly increasing and exhibits PRS, Proposition 2(ii) implies that the optimum is attained by an interval partition. Fix $\theta \in (\underline{\theta}, \overline{\theta})$. Observe that pooling and separating G_ϵ are both feasible since $Q(G)$ is weakly increasing. The principal prefers to pool G_ϵ if

$$E_\epsilon[h(x)]Q(G_\epsilon) > E_\epsilon[h(x)Q(\{x\})] \tag{11}$$

where $E_\epsilon[h(x)] := E[h(x) \mid G_\epsilon]$. Using the definition of local returns to scale (9), inequality (11) becomes

$$E_\epsilon[h(x)]R_\theta(\epsilon) > E_\epsilon\left[\big(h(x) - E_\epsilon[h(x)]\big)\big(Q(\{x\}) - E_\epsilon[Q(\{x\})]\big)\right] = \mathrm{Cov}_\epsilon(h(x), Q(\{x\})). \tag{12}$$

Taking Taylor expansions, one can show that

$$E_\epsilon[h(x)] = h(\theta) + \tfrac{1}{2}h'(\theta)\epsilon + O(\epsilon^2) \tag{13}$$

$$h(x) = h(\theta) + h'(\theta)(x - \theta) + O((x - \theta)^2). \tag{14}$$

Substituting (13) and (14) into the definition of $\mathrm{Cov}_\epsilon(h(x), Q(\{x\}))$ yields

$$\begin{aligned}
\mathrm{Cov}_\epsilon(h(x), Q(\{x\})) &= h'(\theta)Q'(\{\theta\})E_\epsilon\left[\left(x - \theta - \tfrac{1}{2}\epsilon\right)^2\right] + O(\epsilon^3) \\
&= h'(\theta)Q'(\{\theta\})\left[E_\epsilon[(x-\theta)^2] - \epsilon E_\epsilon[x - \theta] + \tfrac{1}{4}\epsilon^2\right] + O(\epsilon^3) \\
&= h'(\theta)Q'(\{\theta\})\left[\tfrac{1}{3}\epsilon^2 - \tfrac{1}{2}\epsilon^2 + \tfrac{1}{4}\epsilon^2\right] + O(\epsilon^3) \\
&= \tfrac{1}{12}h'(\theta)Q'(\{\theta\})\epsilon^2 + O(\epsilon^3)
\end{aligned} \tag{15}$$

where the second line comes from multiplying out the squared term, and the third line comes from taking Taylor expansions of the two conditional expectations. Using (12) and (15), the principal wishes to pool if

$$[h(\theta) + O(\epsilon)]R_\theta(\epsilon) > \tfrac{1}{12}h'(\theta)Q'(\{\theta\})\epsilon^2 + O(\epsilon^3).$$

This is satisfied if (10) holds. □

Proposition 3 provides a sufficient condition for the optimal group structure to exhibit pooling. This result also reveals that the incentive to separate is determined by the relative slope of the objective function, $h'(\theta)/h(\theta)$. Since (MON) says the relative slope of $MR(\theta)$ is higher than the relative slope of θ, Proposition 3 provides another way of looking at the segregation effect.

5.4 *Group quality functions*

We now apply our results to different quality functions, many of which have been used in previous papers. Examples A–F all satisfy either positive or negative returns to scale.

A. *Average quality* Suppose $Q(G) = E[\theta \mid G]$, so the quality of a group is given by the average type of its members. This is used by status papers such as Rayo (2003), Dubey and Geanakoplos (2004), and Moldovanu et al. (2007), as well as matching papers such as Damiano and Li (2007, 2008). The average quality function is particularly attractive since it satisfies both positive and negative returns to scale. As shown by Rayo (2003), one can then analyze objective functions that are non-monotone. In particular, when the ironed $h(\theta)$ is increasing, the principal chooses full separation; when the ironed $h(\theta)$ is constant, the principal chooses full pooling.[11] Thus there is always full separation under welfare-maximization, but there may be regions of pooling under profit-maximization, if $MR(\theta)$ is badly behaved. This suggests welfare-maximization leads to smaller groups than profit-maximization. In comparison, Proposition 1 says that when we allow for different quality functions, the reverse is likely to be true.

B. *Generalized average quality* Suppose $Q(G) = \phi(E[\theta \mid G])$. If $\phi(\cdot)$ is concave and increasing, as suggested by the empirical analysis of Henderson et al. (1978), then the quality function has positive returns to scale.[12] As a result, the welfare- and profit-maximizing group structures are connected (Proposition 2(ii)). The local returns to scale are then given by,[13]

$$R_\theta(\epsilon) = \phi(E[x \mid G_\epsilon]) - E[\phi(x) \mid G_\epsilon] = -\tfrac{1}{24}\phi''(\theta)\epsilon^2 + O(\epsilon^3). \tag{16}$$

Applying equation (10) from Proposition 3, pooling is desirable if

$$-h(\theta)\phi''(\theta) > 2h'(\theta)\phi'(\theta). \tag{17}$$

As a special case, this analysis applies to the Cobb–Douglas quality function, $Q(G) = E[\theta \mid G]^\beta$ with $\beta \in (0,1)$, which is used by Epple and Romano (1998), Nechyba (2000), Caucutt (2002), and the latter parts of Arnott and Rowse (1987). However, despite exhibiting PRS, this quality function does not satisfy equation

[11] Definition: The *ironed version of a function* $h(\theta)$ is defined such that the integral equals the greatest convex minorant of the integral of $h(\theta)$. See Myerson (1981).

[12] Proof: Let $\mathcal{G}_F \succcurlyeq \mathcal{G}_C$ and $\psi = E[\theta \mid \mathcal{G}_F]$. By Jensen's inequality, $\phi(E[\psi \mid \mathcal{G}_C]) \geq E[\phi(\psi) \mid \mathcal{G}_C]$. Taking expectations over θ and applying the law of iterated expectations yields the result.

[13] Equation (16) is attained by using the Taylor expansions

$$E[\phi(x) \mid G_\epsilon] = \phi(\theta) + \tfrac{1}{2}\phi'(\theta)\epsilon + \left[\tfrac{1}{6}\phi''(\theta) + \tfrac{1}{12}\tfrac{f'(\theta)}{f(\theta)}\phi'(\theta)\right]\epsilon^2 + O(\epsilon^3)$$

$$\phi(E[x \mid G_\epsilon]) = \phi(\theta) + \tfrac{1}{2}\phi'(\theta)\epsilon + \left[\tfrac{1}{8}\phi''(\theta) + \tfrac{1}{12}\tfrac{f'(\theta)}{f(\theta)}\phi'(\theta)\right]\epsilon^2 + O(\epsilon^3).$$

(17) when $h(\theta)=\theta$. Indeed, full separation is welfare maximizing and, by Proposition 1, is also profit-maximizing.[14]

C. *Maxmin quality* Suppose $Q(G)=\beta\sup(G)+\alpha\inf(G)$, with $\alpha,\beta\geq 0$. One special case of this is min-quality, $Q(G)=\inf(G)$, where the group is only as good as its worst member. Another special case is max-quality, $Q(G)=\sup(G)$, where the best agent becomes the "leader" of the group. First, if $\mathcal{G}\in\Gamma$, then $\mathcal{Q}(\theta,\mathcal{G})=\mathcal{Q}(\theta,\mathcal{I}(\mathcal{G}))$, so we can restrict ourselves to connected groups.[15] Next, suppose $\theta\sim U[\underline{\theta},\overline{\theta}]$. If $\beta\leq\alpha$ (e.g. min-quality) this exhibits negative returns to scale for $\mathcal{G}\in\Gamma$ and welfare- and profit-maximization entail full separation (Proposition 2(i)). Conversely, if $\beta>\alpha$ (e.g. max-quality) this exhibits positive returns to scale for $\mathcal{G}\in\Gamma$. Local returns to scale are then given by

$$R_\theta(\epsilon)=\tfrac{1}{2}(\beta-\alpha)\epsilon.$$

Since $R'_\theta(0)>0$, the optimal group structure induces pooling (Proposition 3). See Figure 3 for an illustration.

D. *Multiplicative quality* Suppose $Q(G)=\exp\big(-\alpha\int_G(1-\theta)\,dF(\theta)\big)$, where $\alpha>0$. This is a continuous analogue of the production functions in Kremer (1993) and Lazear (2001), where each child is quiet with probability $\theta_i\in[0,1]$ and a class of N students is attentive with probability $\Pi_{i=1}^N\theta_i$ (see Section B.1). In this case, $Q(G)$ exhibits negative returns to scale and full separation is optimal (Proposition 2(i)).

E. *Average quality with multiplicative size effects* Suppose $Q(G)=\phi(E[\mathbf{1}_G])E[\theta\mid G]$. The slope of $\phi(\cdot)$ represents the importance of scale effects. Assuming $\phi(\cdot)$ is increasing, as in Farrell and Scotchmer (1988), this quality function exhibits positive returns to scale. As a result, the welfare- and profit-maximizing group structures are connected (Proposition 2(ii)). The local returns to scale are then given by

$$\begin{aligned}R_\theta(\epsilon)&=E[x\mid G_\epsilon]\left[\phi(F(\theta+\epsilon)-F(\theta))-\phi(0)\right]\\&=\theta\phi'(0)f(\theta)\epsilon+\tfrac{1}{2}\left[\theta\phi'(0)f'(\theta)+\theta\phi''(0)f(\theta)^2+\phi'(0)f(\theta)\right]\epsilon^2+O(\epsilon^3).\end{aligned}$$

The optimal group structure therefore induces pooling if $\phi'(0)>0$ (Proposition 3).

F. *Average quality with additive size effects* Suppose $Q(G)=E[\theta\mid G]+\phi(E[\mathbf{1}_G])$. If $\phi(\cdot)$ is decreasing, then agents crowd each other out, the quality function exhibits NRS, and full separation is optimal (Proposition 2(i)). If $\phi(\cdot)$ is increasing, then

[14] Proof that full separation is welfare maximizing: Since $Q(G)$ satisfies PRS, the optimal group structure is connected. Pick an arbitrary group $G=[a,b]\subset[\underline{\theta},\overline{\theta}]$. We show that welfare in increased by separating G. First, observe that separation is feasible since $Q(G)$ is weakly increasing. Welfare under pooling, conditional on being in G, is $E[\theta\mid G]^{1+\beta}$. Welfare under separation, conditional on being in G, is $E[\theta^{1+\beta}\mid G]$. Since $\beta>0$, Jensen's inequality yields the result.

[15] Proof: Pick $I\in\mathcal{I}(\mathcal{G})$. There exists $G\subset I$ such that $\sup(G)=\sup(I)$. Suppose, by contradiction, that $Q(I)\neq Q(G)$. Then $\alpha>0$ and $\inf(G)>\inf(I)$. There exists $G'\subset I$ such that $\inf(G')=\inf(I)$. Since $\alpha>0$, $\sup(G)\geq\sup(G')$ and $\inf(G)>\inf(G')$, so $Q(G)>Q(G')$, contradicting monotonicity.

there are network effects, the quality function exhibits PRS, and the optimal group structure is connected (Proposition 2(ii)). The local returns to scale are given by

$$\begin{aligned} R_\theta(\epsilon) &= \phi(F(\theta+\epsilon)-F(\theta))-\phi(0) \\ &= \phi'(0)f(\theta)\epsilon+\tfrac{1}{2}\left[\phi'(0)f'(\theta)+\phi''(0)f(\theta)^2\right]\epsilon^2+O(\epsilon^3). \end{aligned}$$

The optimal group structure therefore induces pooling if $\phi'(0)>0$ (Proposition 3).

Example G fails to satisfy either PRS or NRS. It shows that the optimal group structure may not be connected.

G. *Max quality with additive size effects* Suppose the quality of the group depends upon its leader and the number of followers. In particular,

$$Q(G)=\begin{cases}\sup(G) & \text{if } E[\mathbf{1}_G]\le 1/2\\ \sup(G)-1 & \text{if } E[\mathbf{1}_G]>1/2.\end{cases}$$

Here, groups do not take the form of intervals: it is optimal to have two groups, both with mass $1/2$ and containing a very good leader. Since groups overlap, Lemma 2(i) implies that they must have the same quality. For example, if $\theta\sim U[0,1]$, then it is optimal to set $G_1=[1/2,1)$ and $G_2=[0,1/2)\cup\{1\}$.

6. The exclusion effect

In Section 5 we examine the optimal way to segregate different types of agents when the principal serves all agents. In this section we extend the analysis to allow for exclusion. In the education example, these excluded agents may attend a public school or, in the case of universities, enter the workplace.

6.1 *Principal's problem*

There are two possible reasons to exclude an agent. First, the principal might wish to exclude θ if $h(\theta)<0$. Second, the principal can exclude groups to 'monotonize' a non-monotonic quality function, expanding the set of implementable group-structures. To illustrate this latter effect, consider the two-type model in Section 2 and suppose $Q_L>Q_H>Q_{LH}$. If Q_{LH} is sufficiently low, the principal may prefer to separate rather than pool. Since $Q_L>Q_H$, separation is feasible only if the low types are excluded. To simplify the presentation we abstract from this effect by assuming that $Q(G)$ is weakly increasing in G. The working paper version of the paper, Board (2007), extends Propositions 4 and 5 to the more general case.

An agent has an outside option of zero. Given a group structure $\mathcal{G}$, suppose the principal excludes $A\subset[\underline{\theta},\overline{\theta}]$, made up of groups from $\mathcal{G}$. Agents' rents can then be characterized by Lemma 1, where the quality function is given by $\mathcal{Q}(\theta,\mathcal{G})\mathbf{1}_{\neg A}$ and $\neg A:=\{\theta:\theta\notin A\}$. Formally, the principal's problem is to choose a group structure $\mathcal{G}$ and a set of excluded agents A to maximize

$$H=E\left[h(\theta)\,\mathcal{Q}(\theta,\mathcal{G})\mathbf{1}_{\neg A}\right] \tag{18}$$

subject to $\mathcal{Q}(\theta,\mathcal{G})\mathbf{1}_{\neg A}$ increasing in θ.

DEFINITION 2. An allocation rule $\langle \mathcal{G}, A\rangle$ is *nice* if it satisfies

(i) $\mathcal{G} \in \Gamma$

(ii) A is a decreasing set[16]

(iii) A is measurable with respect to $\sigma(\mathcal{I}(\mathcal{G}))$.[17]

The principal's problem is equivalent to choosing a nice allocation rule to maximize (18). To see this, we make two observations. First, any nice allocation rule satisfies monotonicity. Second, any allocation rule $\langle \tilde{\mathcal{G}}, \tilde{A}\rangle$ that satisfies monotonicity can be replaced by a payoff-equivalent nice allocation rule $\langle \mathcal{G}, A\rangle$.[18]

Lemma 4 characterizes the optimal excluded set. This result applies to the welfare-maximization problem and, under (MON), to the profit-maximization problem. Denote the positive and negative components of a function by $\phi(x)^+ := \max\{\phi(x), 0\}$ and $\phi(x)^- := -\min\{\phi(x), 0\}$.

LEMMA 4. *Fix $\mathcal{G} \in \Gamma$. Suppose $h(\theta)$ is quasi-increasing and $Q(G)$ is weakly increasing in G. Then the principal's maximal payoff is given by*

$$H(\mathcal{G}) = E\big[E[h(\theta) \mid I(\mathcal{G})]^+ Q(\theta, \mathcal{G})\big]. \tag{19}$$

PROOF. Fix $\mathcal{G} \in \Gamma$. We seek to choose A to maximize the principal's payoff (18) subject to A being decreasing and $\sigma(\mathcal{I}(\mathcal{G}))$-measurable. We can rewrite (18) as

$$E\big[h(\theta)\mathcal{Q}(\theta, \mathcal{G})\mathbf{1}_{\neg A}\big] = E\big[E[h(\theta)\mathcal{Q}(\theta, \mathcal{G})\mathbf{1}_{\neg A} \mid I(\mathcal{G})]\big] = E\big[E[h(\theta) \mid I(\mathcal{G})]\, \mathcal{Q}(\theta, \mathcal{G})\mathbf{1}_{\neg A}\big].$$

The first equality uses the law of iterated expectations and the second uses the fact that $\mathcal{Q}(\theta, \mathcal{G})$ and A are measurable with respect to $\sigma(\mathcal{I}(\mathcal{G}))$. Pointwise maximization implies that the principal's payoff is maximized when she excludes the set

$$A^* = \{\theta : E[h(\theta) \mid I(\mathcal{G})] < 0\}. \tag{20}$$

Observe that A^* is $\sigma(\mathcal{I}(\mathcal{G}))$-measurable and, since $h(\theta)$ is quasi-increasing, is also decreasing. This yields equation (19). □

6.2 *Welfare- and profit-maximization*

The principal's problem is to choose $\mathcal{G}$ to maximize (19). The excluded set is then given by (20). Proposition 4 shows a profit-maximizing principal excludes too many agents.

[16] Definition: A is *decreasing* if $\theta_H \in A$ and $\theta_H \geq \theta_L$ implies $\theta_L \in A$.

[17] Notation: $\sigma(\mathcal{I})$ is the sigma-algebra generated by $\mathcal{I}$.

[18] Proof: Suppose $Q(\theta, \tilde{\mathcal{G}})\mathbf{1}_{\neg \tilde{A}}$ is increasing. Clearly $\tilde{A}$ is decreasing and is measurable with respect to $\sigma(\mathcal{I}(\tilde{\mathcal{G}}))$. Form a new group structure, $\mathcal{G}$, by pooling all excluded agents into one group and let $A = \tilde{A}$. Since $Q(G)$ is weakly increasing, $\mathcal{G} \in \Gamma$. Hence the new allocation rule $\langle \mathcal{G}, A, \rangle$ is nice and is payoff equivalent to $\langle \tilde{\mathcal{G}}, \tilde{A}\rangle$, as required.

PROPOSITION 4 (Exclusion effect). *Suppose that (MON) holds and $Q(G)$ is weakly increasing in G. Then fewer agents are excluded under welfare-maximization than under profit-maximization (up to sets of measure zero).*

PROOF. Lemma 4 implies that welfare is maximized by $A^W = \emptyset$, so that $A^W \subset A^\Pi$. □

The exclusion effect is analogous to the standard monopoly distortion. Under profit-maximization the principal would like to exclude agents with negative marginal revenue, whereas under welfare-maximization the principal would like to exclude no agents.

Proposition 5 shows that the segregation effect extends to the case where the principal can exclude agents. Notably, this result places no restrictions on the sign of $MR(\theta)$.

PROPOSITION 5 (Segregation effect II). *Suppose (MON) holds and $Q(G)$ is weakly increasing in G. Pick any welfare-maximizing solution, $\mathcal{G}^W$. Then $\Pi(\mathcal{G}^W) \geq \Pi(\mathcal{G})$ for all $\mathcal{G} \in \Gamma$ that are coarser than $\mathcal{G}^W$.*

The proof of this result is given in Section A.3.

Two effects underlie Proposition 5. First, a profit-maximizing principal cares relatively more about high-value agents than a welfare-maximizing principal (see Proposition 1). Second, a profit-maximizing principal is more willing to exclude agents than is a welfare-maximizing principal (see Proposition 4). Hence the smaller group size provides additional flexibility to exclude some agents.

Proposition 6 provides a characterization of the excluded agents. A quality function $Q(G)$ is *increasing* in G if $Q(G_H) \geq Q(G_L)$ whenever G_H is larger than G_L in strict set order.[19]

PROPOSITION 6. *Suppose $h(\theta)$ is increasing and $Q(G)$ is weakly increasing in G.*

(i) It is optimal for the principal to exclude $A^* \subset \{\theta : h(\theta) < 0\}$.

(ii) Suppose that either (a) $Q(G)$ exhibits NRS or (b) $Q(G)$ is increasing in G and exhibits PRS. Then it is optimal for the principal to exclude $A^* = \{\theta : h(\theta) < 0\}$.

PROOF. (i) Let $\overline{D} := \{\theta : h(\theta) < 0\}$. Since $h(\cdot)$ is increasing, $\overline{D}$ is a decreasing set. By contradiction, suppose that $A \setminus \overline{D}$ has positive measure. Form a new group structure $\mathcal{G}'$ by including $A \setminus \overline{D}$ as a single group. Since $Q(G)$ is (weakly) increasing, $\mathcal{G}' \in \Gamma$. Moreover, $\mathcal{G}'$ yields a (weakly) higher payoff than $\mathcal{G}$.

(ii)(a) $Q(G)$ exhibits NRS, so Proposition 2(i) implies that, among all $\mathcal{G} \in \Gamma$, the principal's payoff is maximized by full separation.[20] Using pointwise maximization, profit is maximized when the principal excludes $\overline{D}$.

(ii)(b) $Q(G)$ exhibits PRS, so Proposition 2(ii) implies that, among all $\mathcal{G} \in \Gamma$, the principal's payoff is maximized when groups are intervals.[21] By contradiction, suppose that

[19] Definition: G_H is larger than G_L in *strict set order* if $\min\{\theta, \theta'\} \in G_L$ and $\max\{\theta, \theta'\} \in G_H$ for all $\theta \in G_L$ and $\theta' \in G_H$.

[20] Proposition 2(i) does not allow for exclusion, but the result immediately extends. With exclusion, smaller groups provide more flexibility and, via Jensen's inequality, further increase the principal's payoff.

[21] Proposition 2(ii) does not allow for exclusion, but the result immediately extends. The proof is identical.

$\overline{D}\setminus A$ has positive measure. Form a new group structure $\mathscr{G}'$ by splitting any group G into $G\cap\overline{D}$ and $G\cap\neg\overline{D}$, and consider excluding $\overline{D}$. Since $Q(G)$ is (weakly) increasing, $\mathscr{G}'\in\Gamma$. Moreover, $Q(G)$ is increasing, so $Q(G\cap\neg\overline{D})\geq Q(G)$. As a result, $\mathscr{G}'$ yields a (weakly) higher payoff than $\mathscr{G}$,

$$E[h(\theta)Q(\theta,\mathscr{G})\mathbf{1}_{\neg A}]\leq E[h(\theta)Q(\theta,\mathscr{G})\mathbf{1}_{\neg\overline{D}}]\leq E[h(\theta)Q(\theta,\mathscr{G}')\mathbf{1}_{\neg\overline{D}}]$$

as required. □

Proposition 6(i) says that an agent should be included if $h(\theta)\geq 0$. It also may be optimal to include some agents with $h(\theta)<0$, if they exert a positive externality on the agents with $h(\theta)>0$. Broadly speaking, Proposition 6(ii) says that this form of externality is ruled out if the quality function is increasing.

7. Costly group formation

The segregation effect (Proposition 1) states that groups are finer under profit-maximization than they are under welfare-maximization. With costly group formation this is countered by the appropriability effect: a profit-maximizing principal cannot capture consumer surplus and may not invest enough in creating groups. Examples 2 and 3 illustrate how the appropriability effect can dominate the segregation effect. Proposition 7 then derives sufficient conditions for the segregation effect to dominate the appropriability effect.

In order to focus on the segregation effect, we suppose the principal cannot exclude any agents. The principal's problem is thus to choose $\mathscr{G}\in\Gamma$ to maximize $H(\mathscr{G})-c(\mathscr{G})$, where $H(\mathscr{G}):=E[h(\theta)\mathscr{Q}(\theta,\mathscr{G})]$ and $c(\mathscr{G})$ is an arbitrary cost function.

Example 2 (Appropriability effect I). Suppose $\theta\sim\text{Par}(\alpha,\beta)$, as in Example 1. If there are N groups, the principal's problem is to choose groups $\{G_i\}_{i=1}^N$ to maximize

$$(1-\alpha^{-1})\sum_{i=1}^{N}Q(G_i)\int_{G_i}\theta\,dF-c(\mathscr{G}). \tag{21}$$

This coincides with the welfare-maximizing problem if $\alpha=\infty$. Suppose that $c(\mathscr{G})$ depends on $\mathscr{G}$ only through the number of groups N, and is increasing in N (e.g. N is the number of teachers). Expression (21) is supermodular in (α,N), so there are more groups under welfare-maximization than under profit-maximization. ◊

Example 3 (Appropriability effect II). Suppose that splitting a group increases everyone's quality (e.g. multiplicative quality) and that $MR(\theta)\geq 0$. Hence $\Delta\mathscr{Q}(\theta)=\mathscr{Q}(\theta,\mathscr{G}_F)-\mathscr{Q}(\theta,\mathscr{G}_C)\geq 0$ $(\forall\theta)$, for $\mathscr{G}_C,\mathscr{G}_F\in\Gamma$ such that $\mathscr{G}_F\succcurlyeq\mathscr{G}_C$. Since $\theta\geq MR(\theta)\geq 0$, we have $E[\theta\Delta\mathscr{Q}(\theta)]\geq E[MR(\theta)\Delta\mathscr{Q}(\theta)]$. That is, whenever a profit-maximizer splits a group, a welfare-maximizer also splits the group. ◊

ASSUMPTION (HR). $[1-F(\theta)]/f(\theta)$ is decreasing in θ.

This hazard rate assumption is stronger than the (MON) assumption used in Sections 5 and 6. Proposition 7 shows that the segregation effect continues to hold when group formation is costly if the distribution of types obeys (HR) and the quality function exhibits PRS.

PROPOSITION 7 (Weak Segregation Effect). *Suppose (HR) holds and $Q(G)$ exhibits positive returns to scale. Pick any welfare-maximizing solution, $\mathcal{G}^W$. Then $\Pi(\mathcal{G}^W) \geq \Pi(\mathcal{G})$ for all $\mathcal{G} \in \Gamma$ that are coarser than $\mathcal{G}^W$.*

PROOF. Suppose $\mathcal{G}^W$ maximizes welfare and fix $\mathcal{G} \preccurlyeq \mathcal{G}^W$ such that $\mathcal{G} \in \Gamma$. Hence $E[\theta \Delta \mathcal{Q}(\theta)] \geq c(\mathcal{G}^W) - c(\mathcal{G})$, where $\Delta \mathcal{Q}(\theta) = \mathcal{Q}(\theta, \mathcal{G}^W) - \mathcal{Q}(\theta, \mathcal{G})$. Let $\mathcal{I}^*$ be the coarsest interval partition on which $\Delta \mathcal{Q}(\theta)$ is quasi-increasing. As in Proposition 2, PRS implies that $E[\Delta \mathcal{Q}(\theta) \mid \mathcal{I}^*] \leq 0$. Since $\Delta \mathcal{Q}(\theta)$ is quasi-increasing on each $I^* \in \mathcal{I}^*$, $E[\mathbf{1}_D \Delta \mathcal{Q}(\theta)] \leq 0$ for any decreasing set D.

For decreasing sets $\{D_i\}$ and positive constants $\{a_i\}$, $i \in \{1, \ldots, m\}$,

$$E\left[\sum_i a_i \mathbf{1}_{D_i} \Delta \mathcal{Q}(\theta)\right] \leq 0.$$

Since (HR) implies that $[1-F(\theta)]/f(\theta)$ is decreasing, we can define $\{D_i\}$ such that $\sum_i a_i \mathbf{1}_{D_i} \to [1-F(\theta)]/f(\theta)$ as $m \to \infty$. Hence

$$E\left[\frac{1-F(\theta)}{f(\theta)} \Delta \mathcal{Q}(\theta)\right] \leq 0. \tag{22}$$

Inequality (22) implies that

$$\Pi(\mathcal{G}^W) - \Pi(\mathcal{G}) = E[MR(\theta) \Delta \mathcal{Q}(\theta)] \geq E[\theta \Delta \mathcal{Q}(\theta)] \geq c(\mathcal{G}^W) - c(\mathcal{G}),$$

as required. □

The appropriability effect states that a profit-maximizing principal cannot capture consumer surplus and may not invest enough in group formation. Under (HR) and PRS, consumer surplus is maximized by complete pooling, so a profit-maximizer is willing to invest more in group formation than a welfare-maximizer.

Proposition 7 is more restrictive than the original segregation effect (Proposition 1). First, it assumes that the distribution of types satisfies (HR) rather than (MON). Defining $h(\theta, 0) = \theta$ and $h(\theta, 1) = MR(\theta)$, (MON) implies that $h(\theta, t)$ is log-supermodular, while the stronger (HR) assumption is required for $h(\theta, t)$ to be supermodular. Second, the result assumes that $Q(G)$ satisfies PRS, overcoming the problem in Example 3.

Example 4 provides a tractable numerical illustration of Proposition 7. Observe that Example 4 exhibits zero returns to scale, so the conditions of Proposition 7 are stronger than necessary.

EXAMPLE 4 (Average quality). Suppose $Q(G) = E[\theta \mid G]$ and $c(\mathscr{G})$ depends only on $\mathscr{G}$ through the number of groups, N. By Proposition 2(ii), the optimal group structure consists of intervals. The principal then chooses cutoffs $\{\theta_i\}_{i=0}^N$ to maximize welfare or profit (7). When $\theta \sim U[\underline{\theta}, \overline{\theta}]$, the FOCs for $\{\theta_i\}_{i=1}^{N-1}$ reduce to $(\theta_{i+1} - \theta_i) = (\theta_i - \theta_{i-1})$ under *both* welfare- and profit-maximization. If exclusion is not feasible, then marginal welfare from an extra group is $dW/dN = (\overline{\theta} - \underline{\theta})^2/6N^3$, while the marginal profit from an extra group is $d\Pi/dN = (\overline{\theta} - \underline{\theta})^2/3N^3$. Since $d\Pi/dN \geq dW/dN$, a profit-maximizing principal chooses to have more groups (see Section B.2).

This example shows that, once again, profit-maximization exhibits excessive segregation. However, conditional on choosing the same number of groups, welfare- and profit-maximizing principals choose to divide agents in the same manner. This makes regulation relatively easy: the government needs to restrict only the total number of tariffs; the principal then chooses the welfare-maximizing group structure.[22] ◊

8. TYPE SHIFTS AND GROUP SIZE

In this section we investigate how the optimal group structure changes with an increase in all agents' types. In the education market, these results help us assess how class composition varies (a) with ability and (b) over time.

Group size may change for two reasons. First, as agents' types increase, the returns to scale of the peer technology may change. Second, as agents' types increase, the shape of the objective function changes. In this subsection we consider the technological effect, allowing us to categorize Examples A–G. Subsequently, we adopt a linearity assumption that parses out this technological effect. In Section 8.1, we then show that, under costless group formation, the increase in agents' types leads groups to become coarser. In Section 8.2, we show this result extends to costly group formation if the quality function exhibits positive returns to scale, but reverses under negative returns.

Consider the following experiment. Suppose that types are initially distributed according to $\theta \sim f(\theta)$ on $[\underline{\theta}, \overline{\theta}]$. We examine the effect of an upward shift in the distribution so that $\theta \sim f(\theta - t)$ on $[\underline{\theta} + t, \overline{\theta} + t]$ for $t > 0$. We then compare the size of the group containing θ in the initial distribution to that containing $\theta + t$ in the shifted distribution. As in Section 5, the returns to scale are defined by

$$R(\mathscr{G}_C, \mathscr{G}_F, t) := E[\mathscr{Q}(\theta + t, \mathscr{G}_C + t) - \mathscr{Q}(\theta + t, \mathscr{G}_F + t)].$$

DEFINITION 3. A quality function exhibits

(i) *decreasing returns to scale* (DRS) if $R(\mathscr{G}_C, \mathscr{G}_F, t)$ is decreasing in t for all $\mathscr{G}_F \succcurlyeq \mathscr{G}_C$

(ii) *increasing returns to scale* (IRS) if $R(\mathscr{G}_C, \mathscr{G}_F, t)$ is increasing in t for all $\mathscr{G}_F \succcurlyeq \mathscr{G}_C$

(iii) *constant returns to scale* (CRS) if $R(\mathscr{G}_C, \mathscr{G}_F, t)$ is constant in t for all $\mathscr{G}_F \succcurlyeq \mathscr{G}_C$.

[22] This result depends on the uniform distribution. Suppose we consider splitting $[a, c] \subset [\underline{\theta}, \overline{\theta}]$ into two groups with cutoff b. If $MR(\theta)$ is increasing and convex then, conditional on being in $[a, c]$, one can show that $W(b_H) \geq W(b_L)$ implies $\Pi(b_H) \geq \Pi(b_L)$ for $b_H > b_L$.

In Propositions 2 and 3 we showed that under positive (negative) returns to scale there tends to be pooling (separation). It therefore seems natural that under IRS (DRS), groups tend to become coarser (finer) as t increases. Proposition 8 provides a formalisation of this intuition. It considers the problem of choosing $\mathcal{G}$ to maximize average quality,

$$J(\mathcal{G},t) := E[\mathcal{Q}(\theta+t,\mathcal{G}+t)] - c(\mathcal{G}+t),$$

where $c(\mathcal{G}+t)$ is independent of t.[23]

PROPOSITION 8. *Suppose $\mathcal{G}$ is chosen to maximize $J(\mathcal{G},t)$. Pick $t_H > t_L$ and pick any t_L-optimal solution, $\mathcal{G}^L$.*

(i) Suppose there is DRS. Then $J(\mathcal{G}^L,t_H) \geq J(\mathcal{G},t_H)$ for all $\mathcal{G}$ that are coarser than $\mathcal{G}^L$.

(ii) Suppose there is IRS. Then $J(\mathcal{G}^L,t_H) \geq J(\mathcal{G},t_H)$ for all $\mathcal{G}$ that are finer than $\mathcal{G}^L$.

(iii) Suppose there is CRS. Then $J(\mathcal{G}^L,t_H) \geq J(\mathcal{G},t_H)$ for all $\mathcal{G}$.

PROOF. (i) Pick $\mathcal{G}$ such that $\mathcal{G} \preccurlyeq \mathcal{G}^L$. Since $\mathcal{G}^L$ is t_L-optimal, $R(\mathcal{G},\mathcal{G}^L,t_L) \leq c(\mathcal{G}+t_L) - c(\mathcal{G}^L+t_L)$. Since there are decreasing returns to scale, $R(\mathcal{G},\mathcal{G}^L,t_H) \leq c(\mathcal{G}+t_H) - c(\mathcal{G}^L+t_H)$, as required. Parts (ii) and (iii) are identical. □

We now briefly reconsider the examples in Section 5.4. Of these examples, four exhibit CRS, one exhibits DRS, and one exhibits IRS.

A. *Average quality.* This quality function satisfies CRS and, as shown in Section 5.4, exhibits PRS and NRS.

B. *Generalized average quality.* Recall, if $\phi(\cdot)$ is concave, this quality function exhibits PRS. If $\phi'(\cdot)$ is convex the quality function also obeys DRS.[24] Both these conditions are satisfied by the generalized Cobb–Douglas quality function $Q(G) = (E[\theta|G] - \alpha)^\beta$ with $\beta \in (0,1)$. This suggests that an increase in all agents' types makes groups finer. It also suggests that, if the distribution of types is uniform, then higher types are in smaller groups within a given distribution. This intuition is illustrated in Figure 4.

C. *Maxmin quality.* This quality function exhibits CRS. As shown in Section 5.4, if types are distributed uniformly, then the quality function obeys NRS if $\beta \leq \alpha$ and PRS if $\beta \geq \alpha$.

[23] For simplicity, we ignore the monotonicity condition. In general, the set of implementable group structures, $\Gamma(t)$, may vary with t.

[24] Proof: Pick $\mathcal{G}_F, \mathcal{G}_C$ such that $\mathcal{G}_F \succcurlyeq \mathcal{G}_C$. Let $\psi = E[\theta|\mathcal{G}_F]$. Since $\phi'(\cdot)$ is convex, Jensen's inequality implies $E[\phi'(E[\psi+t \mid \mathcal{G}_C])] \leq E[\phi'(\psi+t)]$. Integrating over $t \in [t_L,t_H]$ and rearranging,

$$R(\mathcal{G}_C,\mathcal{G}_F,t_H) = E[\phi(E[\psi+t_H \mid \mathcal{G}_C])] - E[\phi(\psi+t_H)] \leq E[\phi(E[\psi+t_L \mid \mathcal{G}_C])] - E[\phi(\psi+t_L)] = R(\mathcal{G}_C,\mathcal{G}_F,t_L).$$

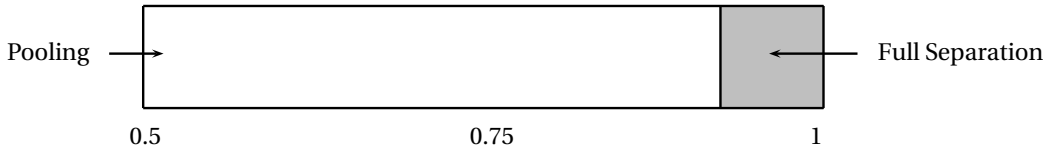

FIGURE 4. Welfare-maximizing group formation: generalized Cobb–Douglas quality. In this figure, $\theta \sim U[0.5,1]$ and $Q(G) = (E[\theta|G] - 0.5)^{0.3}$. Using (17), one can verify that full separation is preferred to small groups when $\theta > (1.3)^{-1} = 0.769$. Using numerical methods, one can show that the welfare-maximizing group structure pools agents below 0.924 into a single group, while those above the cutoff are fully segregated.

D. *Multiplicative quality.* As shown in Section 5.4, this obeys NRS. The quality function satisfies neither IRS nor DRS. However, $R(\mathscr{G}_C, \mathscr{G}_F, t)$ is increasing in t if types are sufficiently close to one. This helps us understand Lazear (2001, Proposition 1) which shows that, with multiplicative quality and homogenous agents, groups are larger when agents' types increase.

E. *Average quality with multiplicative size effects.* If $\phi(\cdot)$ is increasing this quality function exhibits IRS[25] and, as shown in Section 5.4, also exhibits PRS. As a result, groups tend to become coarser as all agents' types increase. This is related to the finding of Farrell and Scotchmer (1988, Proposition 3) that, within a given distribution, group size in the unique stable partition increases in agents' types, if the distribution of types is uniform.

F. *Average quality with additive size effects.* This quality function exhibits CRS. As shown in Section 5.4, this also exhibits NRS if $\phi(\cdot)$ is decreasing and PRS if $\phi(\cdot)$ is increasing.

G. *Max quality with additive size effects.* This exhibits CRS, but satisfies neither PRS nor NRS.

8.1 *Costless group formation*

In this section and the next we adopt the following linearity assumption.

ASSUMPTION (LIN). A vertical shift affects group quality linearly: $Q(G+t) = Q(G) + \lambda t$.

This implies the quality function obeys constant returns to scale, enabling us to parse out the technological effect identified in Proposition 8. Assumption (LIN) also

[25] Proof: Pick $\mathscr{G}_F, \mathscr{G}_C$ such that $\mathscr{G}_F \succcurlyeq \mathscr{G}_C$. Choose $G_C \in \mathscr{G}_C$ and denote the corresponding groups in $\mathscr{G}_F$ by $\{G_i\}$. Suppose G_C has mass m and G_i has mass m_i. Abusing notation, the difference in expected qualities conditional on G_C is $R(t) = \phi(m)E[\theta + t \mid G_C] - \sum_i (m_i/m)\phi(m_i)E[\theta + t \mid G_i]$, which is increasing in t. Summing over $G_C \in \mathscr{G}_C$ yields the result.

implies that the set of implementable group structures,

$$\Gamma(t)=\{\mathscr{G}:\mathscr{Q}(\theta,\mathscr{G}+t)\text{ is increasing in }\theta\in[\underline{\theta}+t,\overline{\theta}+t]\},$$

is independent of t. It is satisfied by Examples A, C, F, and G.

The principal's problem is to choose $\mathscr{G}\in\Gamma$ to maximize[26]

$$H=\int_{\underline{\theta}+t}^{\overline{\theta}+t}h(\theta,t)\mathscr{Q}(\theta,\mathscr{G}+t)f(\theta-t)\,d\theta.$$

Under (LIN) we can change variables to $\tilde{\theta}=\theta-t$. Under welfare-maximization, $h(\theta,t)=\theta$, so the objective becomes $h(\tilde{\theta}+t,t)=\tilde{\theta}+t$. Under profit-maximization, $h(\theta,t)=\theta-[1-F(\theta-t)]/f(\theta-t)$, so the objective becomes $h(\tilde{\theta}+t,t)=MR(\tilde{\theta})+t$. Putting this together, $h(\tilde{\theta}+t,t)=h(\tilde{\theta})+t$. The principal's problem is thus to choose $\mathscr{G}\in\Gamma$ to maximize

$$H(\mathscr{G},t)=\int_{\underline{\theta}}^{\overline{\theta}}[h(\tilde{\theta})+t][Q(\tilde{\theta},\mathscr{G})+\lambda t]f(\tilde{\theta})\,d\tilde{\theta}. \tag{23}$$

Proposition 9. *Suppose $h(\theta)$ is positive and increasing in θ, and that $Q(G)$ satisfies (LIN). Fix $t_H>t_L$ and pick any t_L-optimal solution, $\mathscr{G}^L$. Then $H(\mathscr{G}^L,t_H)\geq H(\mathscr{G},t_H)$ for all $\mathscr{G}\in\Gamma$ that are finer than $\mathscr{G}^L$.*

Proof. The function $h(\theta)+t$ is strictly positive and log-submodular in (θ,t). The rest of the proof is identical to the proof of Proposition 1, although one should use Lemma 5(ii) rather than Lemma 5(i). □

Proposition 9 says that, under either welfare- or profit-maximization, an increase in all agents' types leads groups to become no finer. To understand the result, take a group $[\theta_L+t,\theta_H+t]$ and consider a split that reduces the quality of the low types a lot, while raising the quality of the high types a little. When the agents' types are low (i.e. t is low), the ratio between the highest and lowest types in the group, $(\theta_H+t)/(\theta_L+t)$, is large and this split may increase welfare/profit. Yet when the agents' types are high (i.e. t is high), the ratio between the highest and lowest types in the group is small and the split is less likely to be beneficial.

Our result concerns the group structure as the entire distribution of types shifts. It also suggests that higher types will be in larger groups than lower types within a given distribution if the relative ratio of high types to low types remains constant throughout the distribution (e.g. the density is uniform, ignoring boundary problems). This can be seen in Figure 3, where higher types are in larger groups under both welfare- and profit-maximization.

[26] This assumes the principal cannot exclude.

8.2 *Costly group formation*

With costly group formation, the principal's problem is to choose $\mathscr{G} \in \Gamma$ to maximize $H(\mathscr{G},t) - c(\mathscr{G}+t)$, where $H(\mathscr{G},t)$ is defined by (23) and the cost function $c(\mathscr{G}+t)$ is independent of t.

PROPOSITION 10. *Suppose $h(\theta)+t$ is increasing in θ and that quality satisfies (LIN). Fix $t_H > t_L$ and pick any t_L-optimal solution, $\mathscr{G}^L$.*

(i) Suppose there is PRS. Then $H(\mathscr{G}^L,t_H) \geq H(\mathscr{G},t_H)$ for all $\mathscr{G} \in \Gamma$ that are finer than $\mathscr{G}^L$.

(ii) Suppose there is NRS. Then $H(\mathscr{G}^L,t_H) \geq H(\mathscr{G},t_H)$ for all $\mathscr{G} \in \Gamma$ that are coarser than $\mathscr{G}^L$.

PROOF. (i) Suppose PRS holds. Fix $t_H > t_L$, pick any t_L-optimal solution $\mathscr{G}^L$, and consider $\mathscr{G} \in \Gamma$ such that $\mathscr{G} \succcurlyeq \mathscr{G}^L$. PRS implies that $E[\Delta\mathscr{Q}(\theta)] \geq 0$, where $\Delta\mathscr{Q}(\theta) := \mathscr{Q}(\theta,\mathscr{G}^L) - \mathscr{Q}(\theta,\mathscr{G})$. Using (23), $H(\mathscr{G}^L,t) - H(\mathscr{G},t) = E[(h(\theta)+t)\Delta\mathscr{Q}(\theta)]$ and

$$E[(h(\theta)+t_H)\Delta\mathscr{Q}(\theta)] - E[(h(\theta)+t_L)\Delta\mathscr{Q}(\theta)] = (t_H - t_L)E[\Delta\mathscr{Q}(\theta)] \geq 0.$$

Hence $H(\mathscr{G}^L,t_L) - H(\mathscr{G},t_L) \geq c(\mathscr{G}^L) - c(\mathscr{G})$ implies $H(\mathscr{G}^L,t_H) - H(\mathscr{G},t_H) \geq c(\mathscr{G}^L) - c(\mathscr{G})$, as required. The proof for (ii) is similar. □

Proposition 10 says that (a) under PRS, higher types are in coarser groups, and (b) under NRS, higher types are in finer groups. In comparison, if there is costless group formation then (a) under PRS, higher types are in coarser groups (Proposition 9), and (b) under NRS, there is full separation (Proposition 2). To understand this result, consider the PRS case. Splitting a group has an efficiency effect, reducing the mean group quality, and a distributional effect, benefiting high types while hurting low types. When all types are higher, the importance of the efficiency effect increases while the distributional effect, which depends on the difference between types, remains constant. Hence the principal chooses to create fewer groups.

Proposition 10 considers a shift of the entire distribution of types. Example 5 shows that, under the uniform-maxmin model, a similar result applies within a given distribution of types.

EXAMPLE 5 (Maxmin-quality). Suppose $Q(G) = \alpha \inf(G) + \beta \sup(G)$, $\theta \sim U[\underline{\theta},\overline{\theta}]$, and $c(\mathscr{G})$ depends on $\mathscr{G}$ only through the number of groups, N. As in Section 5.4, the optimal group structure consists of intervals. The welfare-maximizing principal then chooses cutoffs $\{\theta_i\}_{i=0}^N$ to maximize (5). Under zero returns to scale ($\alpha = \beta$), the FOCs for $\{\theta_i\}_{i=1}^{N-1}$ reduce to $(\theta_{i+1} - \theta_i) = (\theta_i - \theta_{i-1})$, as in Example 4, so groups are the same size for all types. Under NRS (i.e. $\alpha \geq \beta$), then $(\theta_{i+1} - \theta_i) \leq (\theta_i - \theta_{i-1})$, so groups are smaller for higher types. Under PRS (i.e. $\alpha \leq \beta$), then $(\theta_{i+1} - \theta_i) \geq (\theta_i - \theta_{i-1})$, so groups are larger for higher types (see Section B.3). ◊

These results have implications for education markets. When considering the optimal classroom size, the assumption of negative returns seems reasonable (although not uncontroversial). Proposition 10 then suggests that more able students should be in smaller classes. Intuitively, when all students become smarter, they have more to gain from a reduction in class size.

When considering the optimal school composition, holding class size constant, Henderson et al. (1978) suggest that positive returns may be the appropriate assumption. Propositions 9 and 10 then suggest that more able students should be in less segregated schools. Intuitively, when all students become smarter, the performance of the least able becomes relatively more important.

9. Conclusion

This paper analyzes how a principal divides agents into groups in the presence of peer effects. With costless group formation, we show that a profit-maximizing principal segregates agents more finely than is socially optimal (the segregation effect) and excludes too many agents (the exclusion effect). We also analyze how the optimal group structure depends upon the returns to scale of peer technology. With costly group formation, we demonstrate that a profit-maximizing firm may not invest enough in group formation (the appropriability effect). However, under positive returns to scale, the segregation effect dominates the appropriability effect.

Our analysis has implications for public policy. Proponents argue that private communities increase welfare by providing safety and comfort for those willing to pay; critics counter that they are discriminatory and isolationist. Our model is consistent with both these arguments, showing how stratification can increase welfare, but also that private provision leads to communities that are insufficiently diverse. This suggests that, in cases where a few local developers have market power, the government should be careful to ensure new developments contain a wide range of housing stock.

This paper also informs the debate on the role of private schools. Much of the discussion over vouchers and public-private partnerships centres on the mantra of parental choice. However, choice is not the aim in itself. This paper has shown that when the options are designed by an organization with market power, then private provision may provide too much choice, introducing excessive segregation. On the positive side, given knowledge of these distortions, there is no reason why an alert regulatory agency cannot mitigate their impact.

We examine a simple model, while allowing for a wide range of peer interactions. A number of extensions are of interest. First, while we consider a single firm, it is important to understand how oligopolistic schools would differentiate themselves. Second, in many practical examples, agents are not atomless and can directly affect the quality of their group. Third, one would like to take account of interactions between the group structure and the quality of the outside option (e.g. the local public school). Fourth, one could allow the principal to choose both inputs (e.g. teachers) as well as group-entry prices. It is hoped that the framework used in this paper can help address some of these issues.

Appendices

A. Omitted proofs

A.1 *Monotone comparative statics*

The following result is a version of Karlin and Rubin (1956, Lemma 1). The proof is essentially identical.

Lemma 5. *Suppose $\Delta\mathscr{Q}(\theta)$ is quasi-increasing on an interval I. In addition, suppose either*

(i) $h(\theta,t)$ is positive, log-supermodular in (θ,t) and decreasing in t, or

(ii) $h(\theta,t)$ is strictly positive and log-supermodular in (θ,t).

Then $E[h(\theta,t)\Delta\mathscr{Q}(\theta)\mid I]$ is quasi-increasing in t.

Proof. (i) By way of contradiction, suppose there exists $t_H > t_L$ such that

$$E[h(\theta,t_L)\Delta\mathscr{Q}(\theta)\mid I]\geq 0 \quad \text{and} \quad E[h(\theta,t_H)\Delta\mathscr{Q}(\theta)\mid I]<0. \tag{24}$$

Since $\Delta\mathscr{Q}(\theta)$ is quasi-increasing on I, we can break it up into positive and negative components. That is, $\Delta\mathscr{Q}(\theta)\geq 0$ on some $I^+\in I$ and $\Delta\mathscr{Q}(\theta)<0$ on $I^-:=I\setminus I^+$. Restrict the state space to I and rewrite (24) as

$$E[h(\theta,t_L)\Delta\mathscr{Q}(\theta)^+]\geq E[h(\theta,t_L)\Delta\mathscr{Q}(\theta)^-] \tag{25}$$
$$E[h(\theta,t_H)\Delta\mathscr{Q}(\theta)^+]< E[h(\theta,t_H)\Delta\mathscr{Q}(\theta)^-]. \tag{26}$$

There are two possible cases. First, suppose that the left-hand side of (25) equals zero. Then the right-hand side of (25) is also zero and, since $h(\theta,t)$ is decreasing in t, the left-hand side of (26) is zero. We thus obtain a contradiction. Second, we suppose the left-hand side of (25) is strictly positive. Multiplying (25) and (26),

$$\begin{aligned}&E[h(\theta,t_L)\Delta\mathscr{Q}(\theta)^+]E[h(\theta,t_H)\Delta\mathscr{Q}(\theta)^-]-E[h(\theta,t_H)\Delta\mathscr{Q}(\theta)^+]E[h(\theta,t_L)\Delta\mathscr{Q}(\theta)^-]\\ &\quad=\int_{I^-}\int_{I^+}[h(\theta_H,t_L)h(\theta_L,t_H)-h(\theta_H,t_H)h(\theta_L,t_L)]\Delta\mathscr{Q}(\theta_H)^+\Delta\mathscr{Q}(\theta_L)^-\,dF(\theta_H)dF(\theta_L)\\ &\quad>0.\end{aligned}$$

This last line contradicts the log-supermodularity of $h(\theta,t)$, as required.

(ii) The proof is nearly identical. To prove that the left-hand side of (25) must be strictly positive suppose, by contradiction, that it equals zero. Since $h(\theta,t)>0$, (25) implies $\Delta Q(\theta)=0$ (a.e.), so the left-hand side of (26) is zero and (26) cannot hold. We thus obtain a contradiction. □

A.2 *Proof of Corollary 1*

Suppose $\mathscr{G}^W \preccurlyeq \mathscr{G}^\Pi$. Lemma 3 implies that $\mathscr{I}(\mathscr{G}^W)\preccurlyeq\mathscr{I}(\mathscr{G}^\Pi)$. Let $\mathscr{I}^*$ be the coarsest interval partition such that $\Delta\mathscr{Q}(\theta):=\mathscr{Q}(\theta,\mathscr{G}^\Pi)-\mathscr{Q}(\theta,\mathscr{G}^W)$ is quasi-increasing.

LEMMA 6. $E[\Delta\mathcal{Q}(\theta) \mid \mathscr{I}^*] \leq 0$.

PROOF. As in Proposition 1, we have $E[\theta\Delta\mathcal{Q}(\theta) \mid \mathscr{I}^*] \leq 0$. For any $I^* \in \mathscr{I}^*$, it follows that

$$0 \geq E[\theta\Delta\mathcal{Q}(\theta) \mid I^*] \geq E[\theta \mid I^*]E[\Delta\mathcal{Q}(\theta) \mid I^*],$$

where the second inequality comes from the fact that a quasi-increasing function is positively correlated with an increasing function (e.g. Persico 2000, Lemma 1). □

Fix $I^* \in \mathscr{I}^*$. Denote the distribution function of $\mathcal{Q}(\theta, \mathscr{G}^\Pi)$, conditional on $\theta \in I^*$, by $F_\Pi(q) := E[\mathbf{1}_{\mathcal{Q}(\theta,\mathscr{G}^\Pi)\leq q} \mid I^*]$. Similarly define the distribution function of $\mathcal{Q}(\theta, \mathscr{G}^W)$, conditional on $\theta \in I^*$, by $F_W(q) := E[\mathbf{1}_{\mathcal{Q}(\theta,\mathscr{G}^W)\leq q} \mid I^*]$.

LEMMA 7. *For any $I^* \in \mathscr{I}^*$, $F_W(q) - F_\Pi(q)$ is weakly quasi-increasing.*[27]

PROOF. The functions $\mathcal{Q}(\theta, \mathscr{G}^W)$ and $\mathcal{Q}(\theta, \mathscr{G}^\Pi)$ are increasing; denote their inverses by $Q_W^{-1}(q) := \inf\{\theta : \mathcal{Q}(\theta, \mathscr{G}^W) > q\}$ and $Q_\Pi^{-1}(q) := \inf\{\theta : \mathcal{Q}(\theta, \mathscr{G}^\Pi) > q\}$. The difference $\mathcal{Q}(\theta, \mathscr{G}^\Pi) - \mathcal{Q}(\theta, \mathscr{G}^W)$ is quasi-increasing on I^*, so $Q_W^{-1}(q) - Q_\Pi^{-1}(q)$ is weakly quasi-increasing. The difference between the distribution functions is

$$F_W(q) - F_\Pi(q) = E[\mathbf{1}_{\theta\leq Q_W^{-1}(q)} - \mathbf{1}_{\theta\leq Q_\Pi^{-1}(q)} \mid I^*].$$

Hence $F_W(q) - F_\Pi(q)$ is weakly quasi-increasing. □

For $I^* \in \mathscr{I}^*$, Lemmas 6 and 7 imply that $[\mathcal{Q}(\theta, \mathscr{G}^W) \mid I^*] \geq_{icv} [\mathcal{Q}(\theta, \mathscr{G}^\Pi) \mid I^*]$, where $\geq_{icv}$ denotes the increasing-concave order (Shaked and Shanthikumar 2007, Theorem 4.A.22(b)). The increasing-concave order is closed under mixtures so $\mathcal{Q}(\theta, \mathscr{G}^W) \geq_{icv} \mathcal{Q}(\theta, \mathscr{G}^\Pi)$ (Shaked and Shanthikumar 2007, Theorem 4.A.8(b)).

A.3 *Proof of Proposition 5*

The method of proof is the same as in Proposition 1. Suppose $\mathscr{G}^W$ maximizes welfare and pick $\mathscr{G} \in \Gamma$ such that $\mathscr{G} \preccurlyeq \mathscr{G}^W$. Since $\mathscr{G}^W$ is welfare-maximizing, $E[\theta\Delta\mathcal{Q}(\theta)] \geq 0$, where $\Delta\mathcal{Q}(\theta) := \mathcal{Q}(\theta, \mathscr{G}^W) - \mathcal{Q}(\theta, \mathscr{G})$. Define $\mathscr{I}^*$ to be the coarsest interval partition on which $\Delta\mathcal{Q}(\theta)$ is quasi-increasing. Applying Lemma 3, $\mathscr{I}(\mathscr{G}) \preccurlyeq \mathscr{I}(\mathscr{G}^W)$. Monotonicity thus implies that $\Delta\mathcal{Q}(\theta)$ is increasing on each $I \in \mathscr{I}(\mathscr{G})$, so $\mathscr{I}^* \preccurlyeq \mathscr{I}(\mathscr{G})$.

LEMMA 8. $E[\theta\Delta\mathcal{Q}(\theta) \mid \mathscr{I}^*] \geq 0$.

The proof of this result is the same as the proof of Proposition 1.

LEMMA 9. $E\big[E[MR(\theta) \mid \mathscr{I}(\mathscr{G}^W)]^+\Delta\mathcal{Q}(\theta) \mid \mathscr{I}^*\big] \geq 0$.

PROOF. Let $h(\theta, 0) = E[\theta \mid \mathscr{I}(\mathscr{G}^W)]$ and $h(\theta, 1) = E[MR(\theta) \mid \mathscr{I}(\mathscr{G}^W)]^+$. Lemma 8 implies that

$$E\big[h(\theta, 0)\Delta\mathcal{Q}(\theta) \mid \mathscr{I}^*\big] = E\big[E[\theta\Delta\mathcal{Q}(\theta) \mid \mathscr{I}(\mathscr{G}^W)] \mid \mathscr{I}^*\big] = E\big[\theta\Delta\mathcal{Q}(\theta) \mid \mathscr{I}^*\big] \geq 0,$$

[27] Definition: A function $\phi : \mathbb{R} \to \mathbb{R}$ is *weakly quasi-increasing* if $\phi(x_L) > 0$ implies $\phi(x_H) \geq 0$ for $x_H > x_L$.

where the first equality uses the fact that $\Delta\mathcal{Q}(\theta)$ is measurable with respect to $\mathcal{I}(\mathcal{G}^W)$ and the second uses the law of iterated expectations. As shown below, $h(\theta,t)$ is log-supermodular in (θ,t). In addition, $h(\theta,t)$ is positive and decreasing in t. Lemma 5(i) then yields the result.

We now complete the proof by verifying that $h(\theta,t)$ is log-supermodular in (θ,t). Let $\psi(\theta,0) := E[\theta \mid \mathcal{I}(\mathcal{G}^W)]$ and $\psi(\theta,1) := E[MR(\theta) \mid \mathcal{I}(\mathcal{G}^W)]$. We wish to show that

$$\psi(\theta_L,0)\psi(\theta_H,1)^+ \geq \psi(\theta_H,0)\psi(\theta_L,1)^+ \tag{27}$$

for $\theta_H > \theta_L$.

We make two observations. First, (MON) implies $MR(\theta)$ is quasi-increasing in θ. Integrating over θ, $\psi(\theta,1)$ is quasi-increasing in θ.

Second, from (MON) we know that $\theta_L MR(\theta_H) \geq \theta_H MR(\theta_L)$. Integrating over (θ_H,θ_L), we thus have

$$\psi(\theta_L,0)\psi(\theta_H,1) \geq \psi(\theta_H,0)\psi(\theta_L,1). \tag{28}$$

We now show that (27) holds. First, if $\psi(\theta_L,1) < 0$, then (27) trivially holds. Second, if $\psi(\theta_L,1) \geq 0$, then the first observation implies $\psi(\theta_H,1) \geq 0$. Equation (28) then implies (27), as required. □

Lemma 9 thus implies that

$$\begin{aligned} E\left[E[MR(\theta) \mid \mathcal{I}(\mathcal{G}^W)]^+ Q(\theta,\mathcal{G}^W) \mid \mathcal{I}^*\right] &\geq E\left[E[MR(\theta) \mid \mathcal{I}(\mathcal{G}^W)]^+ Q(\theta,\mathcal{G}) \mid \mathcal{I}^*\right] \\ &\geq E\left[E[MR(\theta) \mid \mathcal{I}(\mathcal{G})]^+ Q(\theta,\mathcal{G}) \mid \mathcal{I}^*\right] \end{aligned} \tag{29}$$

where the second line uses the fact that $\mathcal{I}(\mathcal{G}) \preccurlyeq \mathcal{I}(\mathcal{G}^W)$. Integrating over (29), $\Pi(\mathcal{G}^W) \geq \Pi(\mathcal{G})$, as required.

B. Omitted details for examples

B.1 *Multiplicative technology*

Kremer (1993) and Lazear (2001) consider a group of agents, $G = \{p_1,\ldots,p_{\mu(G)}\}$, where agent i makes a mistake with probability $p_i = 1-\theta_i$. For example, one can think of a project that requires $\mu(G)$ jobs to be completed. The probability the project is completed successfully is then

$$Q(G) = \prod_{i=1}^{\mu(G)} (1-p_i).$$

We now consider a continuous type analogue to this quality function. Suppose agents' types are distributed according to the absolutely continuous measure μ, where $\mu([\underline{\theta},\overline{\theta}]) = \alpha$. Let $f(\theta) = d\mu(\theta)/\alpha$ be the normalized density. The quality of a group $G \subset [\underline{\theta},\overline{\theta}]$ is determined as follows. First, as in the discrete model, suppose that a project requires $\mu(G)$ jobs to be completed. Second, break each job into k equal tasks. Third,

draw k agents independently from G, where each agent makes a mistake with probability p_i/k. Then let each of these agents do one of the k tasks for each of the $\mu(G)$ jobs. The probability the project is completed successfully is then

$$Q_k(G) = \left[\prod_{j=1}^{k}\left(1-\frac{p_i}{k}\right)\right]^{\mu(G)}. \tag{30}$$

LEMMA 10. *As the number of tasks grows ($k\to\infty$),*

$$Q_k(G) \xrightarrow{p} \exp\left(-\int_G (1-\theta)\,d\mu\right) = \exp\left(-\alpha\int_G (1-\theta)\,dF\right).$$

PROOF. Define

$$\Delta_k(p_i) := \frac{\ln(1)-\ln(1-p_i/k)}{p_i/k}. \tag{31}$$

For each k we draw a new set of agents with error probabilities $\{p_i\}_{i=1}^k$, so $\Delta_k(p_i)$ is a triangular array. Taking logs of equation (30),

$$\begin{aligned}
\ln(Q_k(G)) &= \mu(G)\sum_{i=1}^{k}\ln(1-p_i/k) \qquad (32)\\
&= \mu(G)\frac{1}{k}\sum_{i=1}^{k}-p_i\left[\frac{\ln(1)-\ln(1-p_i/k)}{p_i/k}\right]\\
&= \mu(G)\frac{1}{k}\sum_{i=1}^{k}-p_i\,\Delta_k(p_i)\\
&\xrightarrow{p} -\mu(G)\,E[p\mid G]
\end{aligned}$$

where the second line uses $\ln(1)=0$ and the third line uses (31). Observe that for a given p_i,

$$\lim_{k\to\infty}\Delta_k(p_i) = \frac{d}{dx}\ln(x)|_{x=1} = 1$$

so that $E[p\,\Delta_k(p)] \to E[p|G]$. The forth line of (32) then follows from the weak law of large numbers (e.g. Durrett 1996, p. 41). □

B.2 *Derivation of Example 4*

B.2.1 *Welfare-maximizing problem* The principal chooses $\{\theta_i\}_{i=0}^N$ to maximize welfare

$$W = \sum_{i=1}^{N} E[\theta\mid G_i]\int_{\theta_{i-1}}^{\theta_i}\theta\,dF \tag{33}$$

where $\theta_0 \geq \underline{\theta}$ and $\theta_N = \overline{\theta}$. Assume $\theta \sim U[\underline{\theta},\overline{\theta}]$. The derivative with respect to $\{\theta_i\}_{i=1}^{N-1}$ is

$$\frac{dW}{d\theta_i} = \frac{1}{4(\overline{\theta}-\underline{\theta})}(\theta_{i+1}-\theta_{i-1})(\theta_{i+1}+\theta_{i-1}-2\theta_i).$$

At the global optimum, $(\theta_{i+1}-\theta_i)=(\theta_i-\theta_{i-1})$ and $\theta_i=\theta_0+(i/N)(\overline{\theta}-\theta_0)$. Using (33), welfare is

$$W=\frac{1}{N}\left(\frac{\overline{\theta}-\theta_0}{\overline{\theta}-\underline{\theta}}\right)\sum_{i=1}^{N}\left[\theta_0+\frac{2i-1}{2N}(\overline{\theta}-\theta_0)\right]^2$$
$$=\left(\frac{\overline{\theta}-\theta_0}{\overline{\theta}-\underline{\theta}}\right)\left[\overline{\theta}\theta_0+\frac{(\overline{\theta}-\theta_0)^2}{12N^2}(4N^2-1)\right]. \tag{34}$$

When exclusion is not feasible or not desirable, we have $\theta_0=\underline{\theta}$. Differentiating (34),

$$\frac{dW}{dN}=\frac{(\overline{\theta}-\underline{\theta})^2}{6N^3}.$$

B.2.2 *Profit-maximizing problem* The principal chooses $\{\theta_i\}_{i=0}^N$ to maximize profit

$$\Pi=\sum_{i=1}^{N}E[\theta\mid G_i]\int_{\theta_{i-1}}^{\theta_i}MR(\theta)\,dF, \tag{35}$$

where $\theta_0\geq\underline{\theta}$ and $\theta_N=\overline{\theta}$. Assume $\theta\sim U[\underline{\theta},\overline{\theta}]$. The derivative with respect to $\{\theta_i\}_{i=1}^{N-1}$ is

$$\frac{d\Pi}{d\theta_i}=\frac{1}{2(\overline{\theta}-\underline{\theta})}(\theta_{i+1}-\theta_{i-1})(\theta_{i+1}+\theta_{i-1}-2\theta_i).$$

At the global optimum, $(\theta_{i+1}-\theta_i)=(\theta_i-\theta_{i-1})$ and $\theta_i=\theta_0+(i/N)(\overline{\theta}-\theta_0)$. Using (35), profit is

$$\Pi=\frac{1}{N}\left(\frac{\overline{\theta}-\theta_0}{\overline{\theta}-\underline{\theta}}\right)\sum_{i=1}^{N}\left[\theta_0+\frac{2i-1}{2N}(\overline{\theta}-\theta_0)\right]\left[2\theta_0-\overline{\theta}+\frac{2i-1}{N}(\overline{\theta}-\theta_0)\right]$$
$$=\left(\frac{\overline{\theta}-\theta_0}{\overline{\theta}-\underline{\theta}}\right)\left[\frac{\overline{\theta}(3\theta_0-\overline{\theta})}{2}+\frac{(\overline{\theta}-\theta_0)^2}{6N^2}(4N^2-1)\right]. \tag{36}$$

When exclusion is not feasible or not desirable, we have $\theta_0=\underline{\theta}$. Differentiating (36),

$$\frac{d\Pi}{dN}=\frac{(\overline{\theta}-\underline{\theta})^2}{3N^3}.$$

B.3 *Derivation of Example 5*

Welfare is given by

$$W=\sum_{i=1}^{N}(\beta\theta_i+\alpha\theta_{i-1})\int_{\theta_{i-1}}^{\theta_i}\frac{\theta}{\overline{\theta}-\underline{\theta}}\,d\theta.$$

Differentiating with respect to θ_i,

$$(\overline{\theta}-\underline{\theta})\frac{\partial W}{\partial\theta_i}=\frac{\beta}{2}(\theta_i^2-\theta_{i-1}^2)+(\beta\theta_i+\alpha\theta_{i-1})\theta_i+\frac{\alpha}{2}(\theta_{i+1}^2-\theta_i^2)-(\beta\theta_{i+1}+\alpha\theta_i)\theta_i. \tag{37}$$

Differentiating again,

$$(\overline{\theta}-\underline{\theta})\frac{\partial^2 W}{\partial \theta_i^2}=3(\beta-\alpha)\theta_i-(\beta\theta_{i+1}-\alpha\theta_{i-1}). \tag{38}$$

The FOC (37) induces a quadratic equation in terms of θ_i. This yields two real solutions: the smaller solution always satisfies the SOC (38); the larger solution never satisfies the SOC (38). The smaller solution, θ_i^- may not lie in $[\theta_{i-1},\theta_{i+1}]$. Taking the boundaries into account, the solution is given by $\theta_i^*=\max\{\min\{\theta_i^-,\theta_{i+1}\},\theta_{i-1}\}$.

Write the optimal choice of θ_i as a function of β, $\theta_i^*(\beta)$, and define

$$\mu(\beta):=2\theta_i^*(\beta)-(\theta_{i+1}+\theta_{i-1}).$$

From Example 4 we know that $\mu(\alpha)=0$. We now show that $\mu(\beta)$ is strictly decreasing when $\mu(\beta)=0$. As a result, $\mu(\beta)>0$ for $\beta<\alpha$ and $\mu(\beta)<0$ for $\beta>\alpha$, as claimed in Example 5.

Differentiating the first-order condition (37),

$$(\overline{\theta}-\underline{\theta})\frac{\partial^2 W}{\partial \theta_i\,\partial \beta}=\tfrac{3}{2}\theta_i^2-\tfrac{1}{2}\theta_{i-1}^2-\theta_{i+1}\theta_i.$$

When $\mu(\beta)=0$ then $\theta_i^*(\beta)=\theta_i^-(\beta)$ and

$$(\overline{\theta}-\underline{\theta})\frac{\partial^2 W}{\partial \theta_i\,\partial \beta}=-\tfrac{1}{8}(\theta_{i+1}-\theta_{i-1})^2<0.$$

Hence

$$\tfrac{1}{2}\frac{d}{d\beta}\mu(\beta)=\frac{d}{d\beta}\theta_i^*(\beta)=-\frac{\partial^2 W}{\partial \theta_i\,\partial \beta}\Big/\frac{\partial^2 W}{\partial \theta_i^2}<0,$$

as required.

References

Arnott, Richard J. and John Rowse (1987), "Peer group effects and educational attainment." *Journal of Public Economics*, 32, 287–305. [92, 105]

Benabou, Roland (1993), "Workings of a city: Location, education, and production." *Quarterly Journal of Economics*, 108, 619–652. [93]

Board, Simon (2007), "Monopolistic group design with peer effects." Unpublished paper, Department of Economics, UCLA. [101, 107]

Caucutt, Elizabeth M. (2002), "Educational vouchers when there are peer group effects—size matters." *International Economic Review*, 43, 195–222. [93, 105]

Damiano, Ettore and Hao Li (2007), "Price discrimination and efficient matching." *Economic Theory*, 30, 243–263. [92, 105]

Damiano, Ettore and Hao Li (2008), "Competing matchmaking." *Journal of the European Economic Association*, 6, 789–818. [93, 105]

Dubey, Pradeep and John Geanakoplos (2004), "Grading exams: 100, 99, ..., 1 or A, B, C? Incentives in games of status." Discussion Paper 1467, Cowles Foundation, Yale University. [105]

Durrett, Richard (1996), *Probability: Theory and Examples*, second edition. Duxbury Press, Belmont, California. [121]

Ellickson, Bryan, Birgit Grodal, Suzanne Scotchmer, and William R. Zame (1999), "Clubs and the market." *Econometrica*, 67, 1185–1217. [93]

Epple, Dennis and Richard E. Romano (1998), "Competition between private and public schools, vouchers, and peer-group effects." *American Economic Review*, 88, 33–62. [91, 92, 93, 105]

Farrell, Joseph and Suzanne Scotchmer (1988), "Partnerships." *Quarterly Journal of Economics*, 103, 279–297. [93, 106, 114]

Hanushek, Eric A. (1999), "Some findings from an independent investigation of the tennessee star experiment and from other investigations of class size effects." *Education Evaluation and Policy Analysis*, 21, 143–163. [93]

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin (2003), "Does peer ability affect student achievement?" *Journal of Applied Econometrics*, 18, 527–544. [93]

Helsley, Robert W. and William C. Strange (2000), "Social interactions and the institutions of local government." *American Economic Review*, 90, 1477–1490. [92]

Henderson, Vernon, Peter Mieszkowski, and Yvon Sauvageau (1978), "Peer group effects and educational production functions." *Journal of Public Economics*, 10, 97–106. [93, 97, 102, 105, 117]

Karlin, Samuel and Herman Rubin (1956), "The theory of decision procedures for distributions with monotone likelihood ratio." *Annals of Mathematical Statistics*, 27, 272–299. [118]

Kremer, Michael (1993), "The O-ring theory of economic development." *Quarterly Journal of Economics*, 108, 551–575. [106, 120]

Ladd, Helen F. (2002), "School vouchers: A critical view." *Journal of Economic Perspectives*, 16, 3–24. [90]

Lazear, Edward P. (2001), "Educational production." *Quarterly Journal of Economics*, 116, 777–803. [91, 92, 106, 114, 120]

Mas, Alexandre and Enrico Moretti (forthcoming), "Peers at work." *American Economic Review*. [94, 97]

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995), *Microeconomic Theory*. Oxford University Press, New York. [98]

Milgrom, Paul and Ilya Segal (2002), "Envelope theorems for arbitrary choice sets." *Econometrica*, 70, 583–601. [98]

Moldovanu, Benny, Aner Sela, and Xianwen Shi (2007), "Contests for status." *Journal of Political Economy*, 115, 338–363. [92, 105]

Myerson, Roger B. (1981), "Optimal auction design." *Mathematics of Operations Research*, 6, 58–73. [105]

Nechyba, Thomas J. (2000), "Mobility, targeting, and private-school vouchers." *American Economic Review*, 90, 130–146. [93, 105]

Persico, Nicola (2000), "Information acquisition in auctions." *Econometrica*, 68, 135–148. [103, 119]

Rayo, Luis (2003), "Monopolistic signal provision." Unpublished paper. [90, 91, 92, 105]

Sacerdote, Bruce (2001), "Peer effects with random assignment: Results for dartmouth roommates." *Quarterly Journal of Economics*, 116, 681–704. [93]

Scotchmer, Suzanne (2002), "Local public goods and clubs." In *Handbook of Public Economics. Volume 4* (Alan J. Auerbach and Martin Feldstein, eds.), 1997–2042, Elsevier, Amsterdam. [93]

Shaked, Moshe and J. George Shanthikumar (2007), *Stochastic Orders*. Springer, New York. [119]

Zimmerman, David J. (2003), "Peer effects in academic outcomes: Evidence from a natural experiment." *Review of Economics and Statistics*, 85, 9–23. [93]

Submitted 2007-12-2. Final version accepted 2009-1-28. Available online 2009-1-28.