"Bunching" and Heterogeneity: Characterizing the Monopolistic Provision of Excludable Public goods *

Jishnu Das
The World Bank
jdas1@worldbank.org

Asim Ijaz Khwaja
Harvard University
akhwaja@ksg.harvard.edu

October 2006

Abstract

This paper examines the monopolistic provision of an excludable public good under imperfect information. We examine requirements on the structure of consumer preferences in order to obtain simple optimal pricing schemes. Such simple schemes are characterized by a "threshold" value, where all individuals whose willingness to pay is above the threshold receive the *same* positive quantity of the good and those below are fully excluded. This simple characterization of the optimal pricing scheme then allows us to examine the impact of changes in the willingness to pay distribution on the provision of the public good. In economies with two consumer types increasing inequality leads to greater exclusion. In economies with multiple types, results can be obtained for specific distributional changes. Distributional changes that satisfy the monotone likelihood ratio condition lead to an increase in the monopolist's threshold, but can result in either more or less exclusion.

1 Introduction

The private provision of public goods has generated considerable interest in both research and policy forums. A key issue is how the private sector charges for such commodities and the extent to which individuals may be excluded from consumption. We analyze this problem in an economy where consumers differ in their willingness to pay for the good. Based on the optimal price schedule derived, we then study the implications of changes in the distribution of willingness to pay for the provision of the good.

This paper focuses on excludable public goods provided by a monopolist where such a good is characterized by non-rivalry combined with excludability in consumption. A number of commodities such as highways, television channels and radio frequencies satisfy this requirement. An example that has been the subject of recent interest and discussion is the case of commodities that can be purchased as "electronic downloads" from the internet. This includes "online" music and movies as well as journal articles and tutorials or net-courses. With no congestion costs, these commodities can be thought of as excludable

*We thank Abhijit Banerjee, R.K. Das, Hanan Jacoby, Michael Kremer, Nolan Miller, Peter Norman, Dilip Mokherjee, Debraj Ray, Carolina Sánchez, Tara Vishwanath and Jeffrey Williamson for helpful comments.

public goods—consumption by a single consumer does not decrease the amount available for others, but the firm can reduce the amount available for any subscriber (the number of songs, movies or articles that can be downloaded) through specific restrictions. This paper addresses the important issue of optimal pricing schemes as well as the exclusion of individuals depending on the distribution of willingness to pay in the population for such commodities.

The primary concern is the *informational environment* that the firm operates in. Due to the non-rival nature of the good, in any first-best allocation there is no exclusion and every consumer consumes exactly the same amount. In the case of a monopolist, this first-best allocation is achieved when consumer types are observable and contracts can be written as a function of the consumer's willingness to pay. However, in a more realistic setting, while the firm may know the overall distribution of willingness to pay in the population, it is unable to observe the willingness to pay for *each* consumer. An optimal contract would therefore typically consist of a menu of price-quantity pairs that operates as a screening mechanism to differentiate among different types of consumers.

The paper makes two contributions. The first relates to the form of the optimal contract. The optimal contract consists of a *single* price-quantity pair *as long as* the ratio of marginal willingness to pay between any two consumer-types is monotonically increasing (or decreasing). We call this the "ratio-monotonicity" condition for preferences. With this restriction, all individuals whose willingness to pay exceeds a certain "threshold" receive the same amount of the good while those below choose not to consume the public good at all. This characterization provides some justification for the observed simplicity of contracts in the real world (continuing with our previous example, several on-line firms offer only a single price-quantity pair for internet downloads instead of a menu of options specifying the prices and the number of downloads permitted) but also allows for more complicated pricing schemes to develop in instances where the required restrictions on preferences are not satisfied.

These results generalize previous research on the preference dimension, although in a more specific environment. In our setup, the actions of one consumer do not reveal information about another. Thus strategic interactions between the consumers, such as message games, are ruled out. This setting is plausible where there are a large number of consumers but can also hold in small economies. In contrast, Hellwig (2002) and Norman (2003) derive similar "threshold" results for preferences that automatically satisfy the ratio-montonicity condition, but allow for a more general mechanism design problem in the context of a contribution game. Together with their work, this paper thus provides a more complete characterization of optimal contracts in this environment and their relation to underlying preferences in the economy.

Our second contribution links the provision of the public good to changes in the distribution of willingness to pay. This link has remained unexplored in the literature, despite the impact it may have on exclusion of the public good. The simplicity of the threshold result allows us to examine the impact of heterogeneity on the provision of the public good. For models with two consumer types (with high and low willingness to pay) and for multiple types given specific distributions of willingness to pay (the uniform and the normal), increases in inequality lead to greater exclusion. This result is intuitive. The threshold represents a trade-off between the gains in profit obtained from increasing the mass of individuals receiving the good with the losses incurred through an increase in the informational "rent" of consumers who were above the original threshold value. As inequality in the distribution of willingness to pay increases the latter is more likely to dominate, leading to greater exclusion in the optimal contract.

For multiple types and general distributions over the willingness to pay, the threshold price that the monopolist charges increases with inequality if the distributional change satisfies a monotone likelihood ratio condition. However, these changes can lead both to an increase or decrease in the number of individuals receiving the commodity. For distributional changes that do not satisfy the monotone likelihood ratio condition either, such as mean preserving spreads, it is not possible to obtain unambiguous results either on the threshold price or the number of individuals who receive the commodity. This cautions us in drawing premature welfare conclusions. For example,decreases in inequality can lead to decreased access to the public good and an increase in the price charged: Such effects are likely for distributional changes that reduce inequality by shifting the mass of individuals from just above the threshold to below the threshold. The impact of changes in inequality depends critically on the specific manner in which the change happens.

The remainder of the paper is structured as follows: Section 2 reviews the existing literature and Section 3 presents the benchmark perfect information case followed by a proof of the threshold result in an economy with two types of consumers. This proof is followed by a discussion of the result and an examination of it's implications for the relationship between preferences and the form of the optimal contract. We end the section by extending the proof to an arbitrarily large number of agent types. Section 4 discusses comparative static results relating changes in heterogeneity to the provision of the public good and Section 5 concludes.

2 Review of the Literature

The provision of public goods has been discussed by Bergstrom (1986), Warr (1983), Bernheim (1986) and in more recent contributions by Hellwig (2003) and Norman (2004). Another set of results pertaining to the monopolistic provision of excludable public goods has been obtained by Brito and Oakland (1980) and Schmitz (1997). Our paper differs in important ways from both these sets of studies.

The first branch of the literature is interested in the problem of public good provision as a mechanism design problem. The social planner designs a mechanism and the provision of the public good is studied as the equilibrium of a contribution game. In concurrent work to ours Norman (2004) obtains asymptotically efficient "threshold" results for the mechanism design problem. In contrast, this paper examines the problem of monopolistic provision in economies where contracts that specify allocations to one consumer type as a function of the announcements of others are explicitly ruled out. Restricting the environment to such economies allows us to focus on the mapping between preferences and the optimal contract. Thus, the problem studied here closely follows that discussed in Maskin and Riley (1984) and Laffont and Tirole (1993). As shown, the threshold result and the bunching of contracts is sensitive to the preferences that consumers hold.

While our approach is closer to that of Brito and Oakland (1980) and Schmitz (1997), our results are qualitatively different from both. Brito and Oakland (1980) characterize the optimal contract for this problem under the assumption that price-quantity pairs are *different* for every agent-type so that their main proposition is restricted to contracts that satisfy this property. Our paper is based on the insight that non-rivalry in consumption makes the "bunching" of contracts (several consumer-types receiving the same price-quantity pair) *more* likely than the standard differentiating monopolist's problem (Maskin and Riley 1984). However, non-rivalry in itself does not guarantee a threshold result: Our focus then is precisely to provide the required structure on preferences for the optimal contract to be characterized by a threshold.

Schmitz (1997) analyzes a similar problem, but studies allocation rules only for an indivisible excludable public good—the monopolist is restricted in the choice of quantity to a binary set, $Q \in \{0, 1\}$—and hence, only for goods for which the consumers have constant marginal valuations and constant marginal cost of provision. In this paper, we allow for more general functions with regard to the cost and preference structure of firms and consumers in the economy. The results presented here thus generalize those obtained by Schmitz (1997) to the case of a perfectly divisible public good and a flexible form of preferences.

Finally, the focus on changes in the distribution of willingness to pay has not been discussed in the context of monopolistic provision. In the context of a contribution game, results by Bergstrom (1986), Warr (1983) and others derive a neutrality result—the provision of the public good is independent of the distribution of wealth. This neutrality result also holds in our environment when consumer types are verifiable. However, we show that in the imperfect information environment, increases in inequality can lead to greater exclusion of individuals with low willingness to pay.

3 The Model

3.1 Assumptions

Consider a monopolist providing an excludable public good, $E \in [0, E_{\max}]$ and charging a price (or menu of prices) from a set of consumers.[1] The economy is characterized as follows:

1. **Consumer Types and Preferences:**

- Following Maskin and Riley (1984), the i'th consumer-type in our model has quasi-linear preferences given by:

$$U_{v_i}(E, T) = \int_0^E p(x; v_i)dx - T$$

  where E is the quantity of the public good purchased at price T, v_i indexes the willingness to pay for the $i'th$ consumer-type, and $x(p; v_i)$ is the derived demand curve of the i'th consumer type. Hence $p(x; v_i)$ represents the maximum that this type is willing to pay for the x^{th} *marginal* unit of the good.[2]

[1] The upper bound on E is not necessary but convenient. Moreover, one can imagine that it arises naturally from the supply side due to convex costs.

[2] Thus we rule out incomes effects. This is reasonable as long as the proportion spent on the public good is small enough so that variation in the good's price have negligible effects on income. Note however, that the parameter v in the marginal willingness to pay function allows for differences in demand across consumers. Our setup also corresponds to the standard

- We also assume:

  [$A1$] The demand price $p(x; v_i)$ is non-increasing in x and non-negative. Moreover, $p(x; v_i)$ is twice continuously differentiable, strictly increasing in v_i (higher types have a higher willingness to pay), and decreasing in x.

  [$A1'$] Inada Conditions: $p(x, v_i)$ satisfies $\lim_{x\to 0} p(x; v_i) \to \infty$ and $\lim_{x\to\infty} p(x; v_i) \to 0$ (note that these are the usual marginal utility conditions expressed in terms of $p(x; v_i)$).[3]

- There are N consumer-types in the economy with the i'th type indexed by the parameter v_i ($i = 1$ to N). Types differ in their willingness to pay, $p(x; v_i)$, for the good. We assume independence across types i.e. knowing the type of or actions chosen by a consumer does not reveal any information to the monopolist regarding other consumer types. For simplicity, we also assume a unit measure for the total number of consumers in the economy.[4] Thus in the two-type case ($N = 2$), a proportion m of consumers have high willingness to pay and a proportion $1 - m$ have low willingness to pay. In the case of a continuum of consumer types this extends to a probability distribution $F(v)$ (we drop the subscript i as types are now simply indexed by $v \in [0, \overline{v}]$).

1. **Monopolist:** The monopolist has a convex cost function $C(E)$ and chooses a vector of allocations $(E_1, E_2, ...E_N)$ and $(T_1, T_2, ..., T_N)$ (subscripts denote the contracts offered to each consumer-type) such that

   (a) The vectors of allocations and prices satisfy individual rationality constraints for all consumer types i.e.,.$U_i(E_i, T_i) \geq U_i(0, 0)$ $\forall$ i.

   (b) The vectors satisfy incentive compatibility constraints: Individuals prefer their own allocation to that of any other consumer type so that $U_i(E_i, T_i) \geq U_i(E_j, T_j)$ $\forall i, j$.

   (c) The vector of allocations and prices is such that the monopolist maximizes her profits. Thus, $(E_1, E_2, ..., E_N)$, $(T_1, T_2, ...T_N) \in \arg\max \sum_i m_i T_i - C(\max_i\{E_i\})$ where m_i is the fraction (density function in the case of a continuum of types) of consumers of type i in the economy. The second term in the profit expression follows from the non-rival nature of the commodity.

3.2 Solution: Complete Information

We first briefly discuss the benchmark case where consumer types are verifiable so that the monopolist can specify amounts contingent on type. We analyze the case with a continuum of types (the analysis is similar with N types) and obtain an extension of the standard result for a perfectly discriminating monopolist.

consumer optimization problem studied by Brito and Oakland (1980). In their formulation, consumers choose between a private and a public good, say y and E, and are endowed with an income I. Preferences are additively separable and given by the utility function $U_i(y_i, E_i) = g_i(y_i) + h_i(E_i)$ and the budget constraint implies that $y_i = I_i - T_i$ (when the price of y is normalized as 1). This setup is identical to ours provided the individual's outlay on the public good is small relative to his income, or she has quasi-linear preferences.

[3] Under this assumption, threshold contracts become less likely since, at very low quantities, even those with low willingness to pay have a very high marginal benefit of consuming the good. This would suggest that the monopolist may increase profits by providing a strictly positive amount of the good to even these consumer types. Relaxing this assumption therefore does not change the threshold results obtained below.

[4] We can also allow for a discrete number of consumers of each type as long as indepenence is retained.

Lemma 1 *Under perfect information the monopolist fully extracts surplus from each consumer type. All consumers are offered the same amount E^* of the public good, and each consumer type is charged an amount (increasing in the type) $T(v) = \int_0^{E^*} p(x; v)dx$*

Proof. Under perfect information, the monopolist can extract full rent from all consumers and therefore sets the fees of each consumer type to satisfy each individual rationality constraint i.e. $T(v) = \int_0^{E(v)} p(x; v)dx$. Thus the monopolist's problem is to,

$$\max_{E(v)} \int_0^{\overline{v}} [\int_0^{E(v)} p(x; v)dx]f(v)dv - C(\max E(v))$$

where $f(v)$ is the density function over consumer types. From the form of the maximand, we have $E(v) = E^*$ $\forall\ v \in [0, \overline{v}]$. Towards a contradiction, if $\exists\ \widehat{E} < \max(E(v))$ provided to a certain consumer type, providing $\widehat{E} + \varepsilon$ to that type strictly increases profits, since $C(\max E(v)$ remains unchanged and revenue is strictly increasing in E.

Finally, $E^*$is given by the solution to:

$$\int_0^{\overline{v}} p(E^*; v)f(v)dv = C^{'}(E^*)$$

■

3.3 Incomplete Information with Two Consumer Types

Under incomplete information the monopolist does not observe the type of each individual but knows the distribution of types in the population. The problem is to then choose a set of allocations and prices that satisfy the individual rationality and incentive compatibility constraints for each consumer type.

The main result we obtain is that if preferences are restricted by a *separability* or a (weaker) *ratio-monotonicity* condition on the marginal willingness to pay function $p(x; v_i)$, the optimal contract offered by the monopolist takes a particularly simple form. The monopolist chooses a 'threshold' type v^* s.t. for all $v > v^*$, $E = E(v^*)$ and $T = T(v^*)$ and for all $v < v^*$, $E = T = 0$.

To develop some intuition for this result, we first analyze an economy consisting of two types of consumers, indexed by v_r and v_p, and then extend the result to an arbitrarily large number of consumer types. Let a proportion m of consumers have high willingness to pay (for notational simplicity "rich" types) and $(1 - m)$ have low willingness to pay (for simplicity "poor" types) and let the respective demand price functions be given by $p(x; v_r)$ and $p(x; v_p)$ respectively.[5] For an amount, E, of the public good, utility of each type is given by:

$$\begin{aligned} U_r &= \int_0^{E_r} p(x; v_r)dx - T_r \\ U_p &= \int_0^{Ep} p(x; v_p)dx - T_p \end{aligned}$$

[5] Note that since our formulation rules out income effects, the terms rich (poor) simply refers to consumers with a higher (lower) marginal willingness to pay for the public good. As in Maskin and Riley (1984) though, one can allow for wealth effects as long as they enter directly through preferences.

We define assumptions $[A2]$ and $[A2^{'}]$ as follows on the structure of preferences:

$[A2]$ $p(x;v)$ is a separable function so that $p(x;v_i) = p_1(x)p_i(v_i)$ for $i = r, p$. The subscripts clarify that in general the $p_i(v)$ function can differ across consumer types. This assumption is essentially a requirement that the relative (marginal) willingness to pay for the two consumer types is independent of quantity (i.e. the same when both types consume the same quantity).

$[A2^{'}]$ $R(x) = \frac{p(x;v_p)}{p(x;v_r)}$ is a (weakly) monotonic function with the following boundary conditions: if $R^{'}(.) > 0$ then $R(0) > m$ (*or* $R(E_{\max}) < m$) and if instead $R'(.) < 0$ then $R(0) < m$ (*or* $R(E_{\max}) > m$). This condition implies that the ratio of marginal willingness to pay between the two types is either always greater or less than the fraction of higher types (m) in the economy. This comparison has a natural interpretation for the the monopolist—it compares the maximum marginal return obtained if all consumers of both types are to be induced to buy $(1 * p(x;v_p))$ with that if only higher type consumers need to be induced to buy $(m * p(x;v_r))$.

Note that assumption $[A2^{'}]$ is weaker than $[A2]$ in the sense that the former implies that $R(x)$ is a weakly monotonic function (it is in fact independent of x).

Proposition 1 *Under [A1] and either [A2] or [A2$^{'}$] the optimal contract offered by the monopolist is characterized by a unique threshold. That is, either*

$$\begin{aligned} E_r &= E_p = E^* \text{ (pooling) or,} \\ E_r &= E_r^*,\ E_p = 0 \text{ (separation)} \end{aligned}$$

Furthermore, the optimal quantity that the monopolist produces is such that the marginal willingness to pay for the lowest type consuming a positive amount is equal to the ratio of the marginal cost to the total number of consumers consuming the good and the monopolist provides less than the first-best amount of the good.

Proof. The monopolist's problem is to

$$\max_{E_r,E_p} mT_r + (1-m)T_p - C(\max(E_r, E_p)) \text{ s.t.} \tag{1}$$

$$\int_0^{E_r} p(x;v_r)dx - T_r \geq 0 \text{ (Individual Rationality: Rich)} \tag{2}$$

$$\int_0^{E_p} p(x;v_p)dx - T_p \geq 0 \text{ (Individual Rationality: Poor)} \tag{3}$$

$$\int_0^{E_r} p(x;v_r)dx - T_r \geq \int_0^{Ep} p(x;v_r)dx - T_p \text{ (Incentive Compatibility: Rich)} \tag{4}$$

$$\int_0^{E_p} p(x;v_p)dx - T_p \geq \int_0^{E_r} p(x;v_p)dx - T_r \text{ (Incentive Compatibility: Poor)} \tag{5}$$

Using standard arguments, $E_r^* \geq E_p^*$ and the binding constraints are given by (3) and (4). Substituting for T_r and T_p and rearranging gives us the monopolist's modified problem:

$$\max_{E_r,E_p} \{m \int_0^{E_r} p(x;v_r)dx - C(E_r)\} + \{m[\int_0^{E_p} (p(x;v_p) - p(x.v_r))dx] + (1-m) \int_0^{E_p} p(x;v_p)dx\} \tag{6}$$

This expression can be separated into two components. The first, $\{m\{\int_0^{E_r} p(x;v_r)dx - C(E_r)\}$, uniquely determines E_r^* (as long as $E_p^* < E_r^*$) and therefore the total amount of the good offered by the monopolist. The second component, $\{m[\int_0^{E_p}(p(x;v_p) - p(x.v_r))dx] + (1-m)\int_0^{E_p} p(x;v_p)dx\}$ provides an expression for the optimal E_p^*. A sufficient condition for the proposition to hold is that this second component does not attain a local maximum in the interval $[0, E_r^*]$:[6]

$$\frac{\partial\{m[\int_0^{E_p}(p(x;v_p) - p(x.v_r))dx] + (1-m)\int_0^{E_p} p(x;v_p)dx\}}{\partial E_p} \gtrless 0 \ \ \forall E_p \in [0, E_r^*]$$

by Leibnitz's rule, $$m[(p(E_p;v_p) - p(E_p;v_r))] + (1-m)p(E_p;v_p) \gtrless 0 \ \ \forall E_p \in [0, E_r^*]$$

$$\text{or,} \ \ \frac{p(E_p;v_p)}{p(E_p;v_r)} \gtrless m \ \ \forall E_p \in [0, E_r^*] \tag{7}$$

Under $[A2]$, the requirement simplifies to:

$$\frac{p_p(v_p)}{p_r(v_r)} \gtrless m \ \ \forall E_p \in [0, E_r^*] \tag{8}$$

Since $\frac{p_p(v_p)}{p_r(v_r)}$ is independent of E_p, one inequality is always satisfied (barring the unlikely case that this ratio is equal to m). In cases where $\frac{p_p(v_p)}{p_r(v_r)} > m$, the second component of the monopolist's profit is strictly increasing in E_p, and the monopolist will set $E_p^* = E_r^*$. For the reverse inequality, the second component is strictly decreasing and the optimal $E_p^* = 0$.

Under $[A2']$, if $\frac{p(x;v_p)}{p(x;v_r)}$ is weakly monotonic (say weakly decreasing in x), then:

$$\frac{p(E_{\max};v_p)}{p(E_{\max};v_r)} > m \Longrightarrow \frac{p(E_p;v_p)}{p(E_p;v_r)} > m \ \ \forall E_p \tag{9}$$

and (6) is increasing in E_p. Hence the monopolist will set $E_p^* = E_r^*$. Conversely if $\frac{p(\ 0\ ;v_p)}{p(\ 0\ ;v_r)} < m,\ E_p^* = 0$. A similar threshold result arises with $\frac{p(x;v_p)}{p(x;v_r)}$ (weakly) increasing in x. The weaker assumption $[A2^{'}]$ therefore requires an additional boundary condition.

For the second part of the proof there are two alternatives. If $E_p^* = 0$, (6) is equivalent to $\max_{E_r} m\int_0^E p(x;v_r)dx - C(E_r)$ with the corresponding first order condition $p(E_r^*;v_r) = \frac{C'(E_r^*)}{m}$. Alternatively if $E_p = E_r = E^*$, the problem is to $\max_E \ \{\int_0^E p(x;v_p)dx\} - C(E)$ with the corresponding first order condition $p(E^*;v_p) = C'(E^*)$. The claim follows since $p(E^*;v_i)$ is the marginal willingness to pay for type i when she is consuming E^*.

Finally, the maximization problem in the first-best is:

$$\max_{E_{FB}} m\int_0^E p(x;v_r)dx + (1-m)\int_0^E p(x;v_p)dx - C(E)$$

with the first order condition: $$mp(E_{FB}^*;v_r) + (1-m)p(E_{FB}^*;v_p) = C'(E_{FB})$$

If $E_p^* = 0$, the corresponding first order condition in the case of incomplete information is $mp(E_r^*;v_r) = C'(E_r^*)$. Suppose $E_r^* > E_{FB}^*$. By $[A1]$, this implies that $mp(E_{FB}^*, v_r) > mp(E_r^*;v_r)$. But since $(1-m)p(E_{FB}^*, v_p) > 0$, we have $mp(E_{FB}^*, v_r) + (1-m)p(E_{FB}^*, v_p) > mp(E_r^*;v_r) \Longleftrightarrow C'(E_{FB}) > C'(E_r^*)$.

[6] Note that an even weaker condition would be for the second component not to attain a global maximum in the open interval $(0, E_r^*)$ but this does not lend itself well to a condition on preferences.

From the convexity of the cost function, this establishes a contradiction with our assumption $E_r^* > E_{FB}^*$. A similar argument follows for the case where $E_p = E_r$ (using the property $p(E, v_r) > p(E, v_p)$ under [$A1$]), completing the proof. ■

3.4 Discussion

The above result shows that, under certain restrictions on preferences, even when the poor are willing to pay an infinite amount on the margin for their first ε units of consumption (implied by the Inada conditions), they still may not be offered a positive amount of the good. To provide further intuition, Figure 1 represents preferences in (E, T) space. The result on the bunching of contracts suggests the following: Consider a fully separating contract, $(E_r = E_r^*, E_p = 0)$, depicted by A. Suppose the monopolist marginally increases E_p, so that the poor consumer-types optimally choose (dT_p, dE_p). The gain from this contract is an increase in revenue by the amount dT_p. To ensure that the incentive compatibility of the rich is maintained the rich must be given the new contract $(E_r^{'}, T_r^{new})$ with a loss in revenue given by $(T_r - T_r^{new})$. The proof then shows that if the gain is larger than the loss, the inequality is maintained *for all subsequent increases in* E_p till $E_p^* = E_r^*$. Alternatively, if the monopolist suffers a net loss in revenue from this change, she optimally sets $E_p^* = 0$.

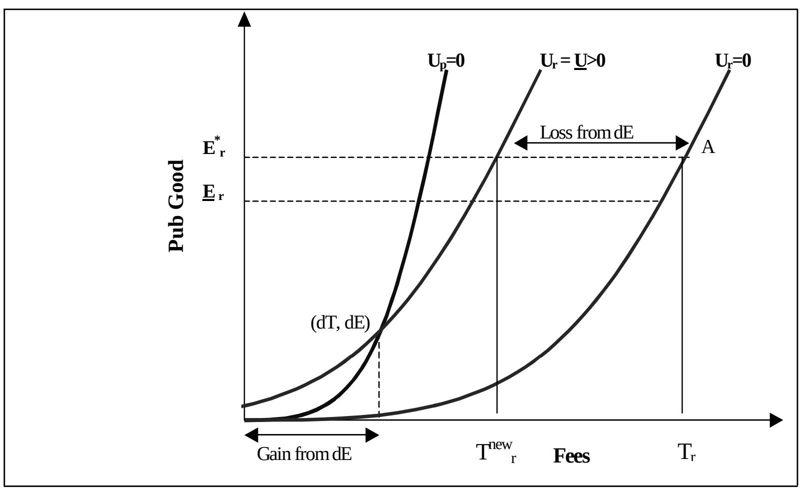


Figure 1: Optimal Contracts with Decreasing Marginal Willingness to Pay

The proof relies on two assumptions: the non-rival nature of the commodity and the form of the utility function, whereby the indifference curves represent horizontal shifts along the (E, T) axis and the ratio of slopes of the two types' indifference curves remains the same for all value of E (under $A2$). The assumption that E is non-rival in consumption implies that we can restrict ourselves to the comparison *only in terms of revenue* as long as E^* is fixed—providing an extra (dE) to the poor consumer has no impact on costs. The role of the second assumption is made clear by recognizing that, since indifference curves are horizontal shifts in (E, T) space, we can decompose the change in profits from moving to a new level of E into two components. The first component is the change in profits from moving *along* the indifference curve of the poor, and the second is the change in profits from the movement *of* the indifference curve of the rich. The second term represents the additional utility (and hence decrease in fees raised) that must be provided to the rich consumer and, if the total amount of the good remains unchanged, this change is equal to the change evaluated at the initial public good level of the poor.

Under the Inada conditions, while it is true that the poor consumer is willing to pay on the margin a very large amount for the first ε units of public good, it is also true that the rich consumer is always willing to pay even more for this first marginal unit. Thus the rich consumer will strictly prefer this new contract to the one he was receiving before (which gave him zero rent). By accepting this new ε public good contract he earns a strictly positive rent and this savings in fees is a function not only of the poor consumers marginal willingness to pay at zero, but also that of the rich consumer's—both the rich and poor consumers' initial marginal conditions are relevant and so the Inada conditions affect both. Thus to be able to offer this new ε public good contract, the monopolist must provide at least as much of an utility increase as the rich gets from accepting the ε contract. In other words, she has to lower fees by the same amount as the savings in fees by the rich on that first marginal unit of public good offered to the poor.

Since the restriction imposed by $[A2]$ (or $[A2']$) is the critical assumption for the optimal contract to be characterized by a threshold, it is useful to restate this condition. Letting $h_i(E) = \int_0^E p(x; v_i)dx$ we can rewrite consumer type i's utility, $U_i = g_i(I - T) + h_i(E)$ where I is his income and $g_i(.)$ utility over a private good (money). In this case assumption $[A2]$ requires that preferences can be obtained as linear transformations of each other (i.e. $h_i(.) \equiv \beta_j h_j(.) + \alpha_j \quad \forall\, i, j$). Stated in these terms it is not unreasonable to think that consumers may (to an approximation) have such preferences, i.e., the "rich" (those who care more about the public good) are on the margin willing to pay a multiple of what the "poor" are. Similarly, the weaker condition, $[A2']$ holds as long as consumer-types do not show "preference reversals" of the form where at a given consumption level Consumer A has a higher marginal willingness to pay for the commodity than Consumer B, but this inequality reverses at other consumption levels.

3.5 Incomplete Information with Multiple Consumer Types

Proposition 1 continues to hold when we extend our model to a continuum of types.[7] Let consumer type now be indexed by $v \in [0, \overline{v}]$ and the distribution of types be given by the cumulative distribution $F(v)$. As before, $U_v = \int_0^E p(x; v)dx - T$, where T is the price a consumer pays for consuming E units of the public good. Moreover, $p(x; v)$ is strictly increasing in v and non-increasing in x.[8] We will refer to consumers of a "higher type" as those characterized by a higher v and as usual, each consumer type experiences diminishing utility from every additional unit of the public good. To prove the equivalent of Proposition 1 we state the following three claims leading to the threshold result. The proofs are fairly standard and are presented in the appendix. We state these claims for N consumer types (index by v_i) and then extend them to a continuum of types.

Claim 1 *If a consumer of a given type weakly prefers a contract with higher public good to another, any*

[7] We get similar results if we instead look at a large but finite number of types.

[8] As before $p(x; v)$ permits different functional forms for different types i.e.$p_v(x, v)$. For tractability we will often keep the same functional form for the willingness to pay function.

consumer of a higher type will strictly prefer the former to the latter contract.

$$\forall E_2 \geq E_1 \& \quad (E_1, T_1) \neq (E_2, T_2), \quad U_i(E_1, T_1) \leq U_i(E_2, T_2) \Rightarrow U_j(E_1, T_1) \leq U_j(E_2, T_2) \; \forall j > i$$

Claim 2 *Any optimal menu of contracts offered by the monopolist,* $(E_i^*, T_i^*)_{i \in \Lambda}$*, satisfies the condition that the contract chosen by the higher consumer type offers an equal or higher level of public good. i.e.* $E_j^* \geq E_i^* \forall j > i$*, where* (E_i^*, T_i^*) *denotes* i'*s preferred contract choice from the menu of contracts offered.*

Claim 3 *If* $(E_i^*, T_i^*)_{i \in \Lambda}$ *represents the optimal (for the monopolist) implementable menu of contracts, then it must be that for any given consumer type, his downward incentive constraint will always be binding i.e.* $(E_i^*, T_i^*) \sim_i (E_{i-1}^*, T_{i-1}^*) \; \forall i$

Using Claim (3):[9]

$$\begin{aligned} & (E_i^*, T_i^*) \sim_i (E_{i-1}^*, T_{i-1}^*) \\ \Rightarrow \quad & U_i^* = \int_0^{E_i^*} p(x; v_i) dx - T_i^* = \int_0^{E_{i-1}^*} p(x; v_{i-1}) dx - T_{i-1}^* \end{aligned}$$

Thus:

$$\begin{aligned} U_i^* - U_{i-1}^* &= \int_0^{E_{i-1}^*} \left[p(x; v_i) - p(x; v_{i-1}) \right] dx \\ &= \int_{v_{i-1}}^{v_i} \int_0^{E_{i-1}^*} \frac{\partial p(x; y)}{\partial v} dx dy \end{aligned}$$

Since $U_0^* = 0$:

$$U_k^* = \sum_{i=1}^{k} \left[\int_{i-1}^{i} \int_0^{E_{i-1}^*} \frac{\partial p(x; y)}{\partial v} dx dy \right] \tag{10}$$

Taking the limit as consumer types tend to infinity (and denoting consumer types simply by the parameter v now i.e. dropping the subscripts on v) , $U_v^* = \int_0^v \int_0^{E_y^*} \frac{\partial p(x;y)}{\partial v} dx dy$ so that:

$$T_v^* = \int_0^{E_v^*} p(x; v) dx - \int_0^v \int_0^{E_y^*} \frac{\partial p(x; y)}{\partial v} dx dy \tag{11}$$

This expression has a natural interpretation. The first term represents the price a fully discriminating monopolist would charge under perfect information. The second term is the loss in the price charged from the information rent that the monopolist pays to higher consumer types in order to reveal their type. The profit of the monopolist is given by:

$$\int_0^{\overline{v}} \left[\int_0^{E_v^*} p(x; v) dx - \int_0^v \int_0^{E_y^*} \frac{\partial p(x; y)}{\partial v} dx dy \right] dF(v) - \widetilde{C}(E_{\overline{v}}) \tag{12}$$

After integration by parts and further simplification, we get:

[9] For notational simplicity we abuse notation slightly by referring to types as v and $v-1$ etc. when referring to v_i and v_{i-1} respectively.

$$\int_0^{\overline{v}} \int_0^{E_v^*} \left[f(v)p(x;v) - (1-F(v)) \frac{\partial p(x;v)}{\partial v} \right] dxdv - \widetilde{C}(E_{\overline{v}}) \tag{13}$$

For each consumer type, the amount E offered by the monopolist represents the following trade-off: By offering (all) consumers of type v a marginal unit of public good the monopolist increases her revenue by $(f(v)p(x;v))$. On the other hand, since type v is now offered more, to satisfy the incentive compatibility of all types higher than v, their informational rent must be increased and their prices reduced. This decline in revenues is given by the second expression $(1-F(v))\frac{\partial p(x;v)}{\partial v}$.[10] Thus the monopolist's problem is to choose functions $E^*(v)$ and $T^*(v)$ that maximize the above expression subject to the constraint that $E^*(v)$ is non-decreasing (recall from above that this constraint has to be satisfied in any equilibrium).

To arrive at the "threshold" result we restate assumptions $[A2]$ and $[A2']$ for a continuum of types:

$[\widetilde{A2}]$ $p(x;v)$ is a separable function that can be written as $p(x;v) = p_1(x)p_v(v)$.[11]

OR

$[\widetilde{A2'}]$ $R(x,v) = \frac{p(x;v)}{\frac{\partial p(x;v)}{\partial v}}$ is (weakly) monotonic in x with additional boundary conditions: If $R(.)$ is increasing the boundary conditions are $\forall v$ either $R(\ 0\ ,v) > \frac{1-F(v)}{f(v)}$ *or* $R(E_{\max},v) < \frac{1-F(v)}{f(v)}$ and vice versa if $R(.)$ is decreasing in x.

Proposition 2 *If either $[\widetilde{A2}]$ or $[\widetilde{A2'}]$ holds the monopolist's optimum is characterized by a simple threshold rule: All consumers of type $v \geq \underline{v}$ are offered the same contract $E^*(\underline{v}), T^*(\underline{v})$, whereas all types $v < \underline{v}$ are not provided the good (offered a contract where $E^*(v) = 0$ and therefore $T^*(v) = 0$).*

Proof. The monopolist solves

$$\max_{E(v)} \int_0^{\overline{v}} \int_0^{E_v^*} \left[f(v)p(x;v) - (1-F(v)) \frac{\partial p(x;v)}{\partial v} \right] dxdv - \widetilde{C}(E_{\overline{v}}) \quad s.t. \quad \frac{\partial E^*(v)}{\partial v} \geq 0$$

Condition $[\widetilde{A2'}]$ guarantees that the integrand in the above expression is always either (weakly) positive or negative for a given v and hence, ignoring the non-decreasing constraint on $E^*(v)$, this implies that $E^*(v)$ should be equal to 0 when the integrand is negative and set to it's maximum optimal value, $E(\overline{v})$, when it is positive (since the cost of E only depends on the maximum E provided, it is never optimal to provide different positive amounts of public good). However, such a solution would violate the non-decreasing constraint on $E^*(v)$. Adding this constraint gives our desired threshold result.

This is clearer under $[\widetilde{A2}]$; Substituting $p(x;v_i) = p_1(x)p_v(v)$ the monopolist's problem can be re-written as:

$$\max_{E(v)} \int_0^{\overline{v}} \left[\left\{ f(v)p_v(v) - (1-F(v)) \frac{\partial p_v(v)}{\partial v} \right\} \int_0^{E(v)} p_1(x)dx \right] dv - \widetilde{C}(E_{\overline{v}}) \quad s.t. \quad \frac{\partial E^*(v)}{\partial v} \geq 0$$

Since $p_1(x) \geq 0$, $\int_0^{E(v)} p_1(x)dx \geq 0$ for $E(v) \geq 0$. However, the 'coefficient' of this integrand is either positive or negative for a given type v. Thus as before, we obtain the desired threshold result. Moreover, in this

[10] To compare this profit function to the two type case, we can re-write (6) as $\int_0^{E_p}(1-m)p(x;v_p) - m\{p(x;v_r) - p(x.v_p)\}dx + \int_0^{E_r} mp(x;v_r)dx - C(E_r)$,
The gains and losses are now given by the equivalent expressions $(1-m)p(x;v_p)$ and $m\{p(x;v_r) - p(x.v_p)\}$.

[11] We abuse nottation slightly to allow for different consumer types to have both a different argument in the $p(.)$ function but also a different functional form for this function (i.e. $p_v(.)$).

simpler formulation, the monopolist first chooses $\underline{v}$ s.t.:

$$\underline{v} \in \arg\max_{v} \int_{v}^{\overline{v}} \left\{ f(s)p_v(v) - (1-F(s)) \frac{\partial p_v(s)}{\partial s} \right\} ds$$

and then the amount of public good to offer to all types $v \geq \underline{v}$:

$$\max_{E} \int_{\underline{v}}^{\overline{v}} \left[\left\{ f(s)p_v(v) - (1-F(s)) \frac{\partial p_v(s)}{\partial s} \right\} \int_0^E p_1(x)dx \right] ds - \widetilde{C}(E)$$

All types $v < \underline{v}$ are not offered any of the public good. ■

While the conditions used to obtain the threshold result are analogous to that in the two-type case, $[\widetilde{A2'}]$ is weaker in terms of the structure imposed on preferences. However, the appeal of $[\widetilde{A2'}]$ is lessened due to the boundary condition requirements, and more so since these boundary conditions may hold for a given distribution of types but then be violated for changes in this distribution. This is particularly relevant when considering how the threshold varies with changes in the distribution of consumer types and for the remainder of the paper we will thus impose the stronger $[\widetilde{A2}]$ condition on preferences.

4 Changes in Heterogeneity

The primary focus of this section is to map changes in the distribution of willingness to pay to the monopolist's threshold, ultimately presenting some insights on how heterogeneity in the distribution of willingness to pay may affect the amount and distribution of the public good provided. This is important given that such public goods may be increasingly privately provided in economies with varying degrees of (preference) heterogeneity. This leads to very different welfare implications depending on the degree of exclusion and prices charged by the private provider.

In the case of perfect information a "neutrality result" obtains, whereby the provision of the excludable public good depends only on (and is increasing in) the mean of the willingness-to-pay distribution. In the case of imperfect information, when there are two types increases in inequality always (weakly) increase the price and degree of exclusion. With multiple types the threshold of the monopolist is directly linked to the hazard rate of the distribution. This leads to a general result on the class of distributional changes that satisfy the *monotone likelihood ratio condition*: Any distributional change (for distributions with increasing hazard rates, satisfied for instance by the uniform, exponential or normal) leads to an increase in the monopolist's threshold.[12] However, this can lead to *both an increase or decrease* in the mass of people excluded from the good. More general results (without placing restrictions on the type of the distribution) do not exist; for instance, mean-preserving spreads have an ambiguous effect on *both* the threshold and the mass of individuals that receive the good. This raises a note of caution for both theoretical and empirical work that attempts to analyze welfare implications of increases in (preference) heterogeneity. The simple intuition that decreases in

[12] The use of the monotone likelihood ratio condition was first considered by Milgrom (1981) and more recently by Athey (2002) in examining the comparative static properties of different economic phenomena. As Milgrom (1981) suggests, this condition is satisfied by a large number of distributions and is thus plausible for a large number of empirical applications. Examples include the normal and the exponential (with respect to the mean).

inequality lead to improved access to the public good through lower prices and a smaller mass of individual's excluded may not be true.

4.1 Perfect Information

Claim 4 *(Neutrality Result): Under perfect information the amount of public good provided to each type depends only on the mean of the willingness to pay distribution.*

Proof. Under perfect information the monopolist offers every agent the same amount $E^*$of the public good such that:

$$\begin{aligned} \int_0^{\overline{v}} p_1(E^*)p_v(v)f(v)dv &= \widetilde{C}'(E^*) \\ or,\ p_1(E^*)\int_0^{\overline{v}} p_v(v)f(v)dv &= \widetilde{C}'(E^*) \end{aligned}$$

The neutrality result follows directly from this expression. To see this formally, define the random variable $Z = p_v(v)\int_0^E p_1(x)dx = p_v(v)K(E)$. The density function of Z, $f(z)$ is then given by $f(z) = f(v)\frac{dv}{dz}$ and the mean of the random variable Z is

$$\begin{aligned} & \int_{p_v(0)K(E)}^{p_v(\overline{v})k(E)} zf(z)dz \\ = & \int_{p_v(0)K(E)}^{p_v(\overline{v})k(E)} p_v(v)K(E)f(v)\frac{dv}{dz}dz \\ = & \int_0^{\overline{v}} p_v(v)K(E)f(v)dv \\ = & \ K(E)\int_0^{\overline{v}} p_v(v)f(v)dv \end{aligned}$$

Thus, the optimal amount E^* depends only on the mean of Z. Hence for any two distributions with the same mean, μ, the first order condition is $p_1(E^*)\mu = \widetilde{C}'(E^*)$. Since this depends only on the mean and not the distribution of Z, the proof is complete. ■

The intuition for this neutrality result is straightforward—with non-rivalry in consumption, the monopolist only offers contracts that fully extract the rent from each consumer type. Thus, the total wealth that she can extract is independent of the distribution of wealth as long as mean wealth remains unchanged. This result is a direct extension of previous neutrality results (Bergstrom 1986 and Warr 1983) to the case of excludable public goods provided by a monopolist.[13] As in the literature on the relationship between the wealth distribution and the provision of a public good through contribution games, the central intuition that only the total willingness to pay matters remains unchanged.

[13] Although recent work by Bardhan and others (2002) shows that in the case of impure public goods (public goods where an individual's benefit from the good depends both on its contribution and the overall contribution to the good) the neutrality result breaks down.

4.2 Incomplete Information

In our example with two types of consumers the impact of inequality on the provision of the public good is unambiguous. Returning to Equation (8), we have $E_p^* = E_r^*$ if $\frac{p_p(v_p)}{p_r(v_r)} > m$ and $E_p^* = 0$ if $\frac{p_p(v_p)}{p_r(v_r)} < m$. From this characterization, if *either* the ratio of willingness to pay for the poor compared to the rich decreases, *or* the proportion of rich types in the economy increases, the poor are more likely to be excluded from receiving the good. Since any mean preserving spread in the willingness to pay in the two type case will necessarily decrease the ratio of willingness to pay for the poor compared to the rich, inequality will tend to result in greater exclusion for the poor. This intuition does not, however, generalize to our model with a continuum of types.

Consider a simplified version of our model with $p_1(x) = 1$ and $p_{v_i}(v_i) = v_i$. Therefore v_i is the marginal willingness to pay of consumer type i is distributed with density $f(v_i)$. The willingness to pay for an amount $E = \int_0^E p(x) v_i dx = K(E)v_i$, and $U_i = K(E)v_i - T_i$. We are interested in examining the impact of a change in the distribution of v_i, $f(v_i)$.

Denote θ as the threshold consumer type; from Proposition (2), the monopolist chooses θ such that all consumers with $v_i \geq \theta$ receive the same amount of the good, and those with $v_i < \theta$ receive 0. The total mass of individuals who receive the good is thus $\int_\theta^\infty f(t)dt$. Further, the IR constraint of type θ must bind, so that $T = K(E)\theta$. The monopolist's profit is then $K(E)\theta \int_\theta^\infty f(t)dt - C(E)$. As before, this program can be maximized in two steps: Since $\theta \int_\theta^\infty f(t)dt$ is independent of E, the monopolist first maximizes this term with respect to θ, and then chooses E and T optimally. Our interest is in the first step of the maximization.[14]

We define the following:

Definition 1 *Consider any two random variables X, Y with common support $[a, b]$, associated density functions f_X, g_Y and distribution functions F_X, G_Y. Then, F_X, G_Y satisfy the monotone likelihood ratio condition (MLRC) (written as $F_X \succ_{MLRC} G_Y$) iff $\frac{f(x)}{g(x)} \leq \frac{f(y)}{g(y)} \forall x \leq y$.*

Definition 2 *$F(x, r)$ denotes a family of distributions that satisfy the MLRC condition. We define the index r such that $r' > r \Rightarrow F(x, r') \succ_{MLR} F(x, r)$. i.e. a distribution with a higher r MLRC dominates one with a smaller r.*

Definition 3 *The hazard rate of any random variable X with associated density function $f(x)$ is defined as $h(x) \equiv \frac{f(x)}{1-F(x)}$*

We then have:

Claim 5 *If $F(x, r')$ MLRC dominates $F(x, r)$ then $F(x, r')$ also has a lower hazard rate than $F(x, r)$ at all points of the support. Stated formally, $F(x, r') \succ_{MLR} F(x, r) \Rightarrow h(x, r') < h(x, r) \forall x$*

[14] Note that this expression can be obtained by subsituting for in $p_2(v_i) = v_i$ in equation (3). Doing so gives $\max_{v,E} \int_v^{\overline{v}} \left[\{f(s)s - (1 - F(s))\} \int_0^E p_1(x)dx \right] ds - \widetilde{C}(E)$ or equivalently $\max_{v,E} K(E) \int_v^{\overline{v}} [\{f(s)s - (1 - F(s))\}] ds - \widetilde{C}(E)$.The integrand can then be rewritten as $v \int_v^{\overline{v}} f(s)$ using integration by parts.

Proof. Note that $F(.,r') \succ_{MLR} F(.,r) \Rightarrow f(y,r') \geq \frac{f(x,r')}{f(x,r)} f(y,r) \forall x \leq y$. By definition,

$$\begin{aligned} h(x,r') &= \frac{f(x,r')}{1-F(x,r')} = \frac{f(x,r')}{\int_a^b f(s,r')ds - \int_a^x f(s,r')ds} \\ &= \frac{f(x,r')}{\int_x^b f(s,r')dt} \leq \frac{f(x,r')}{\int_x^b f(s,r)\frac{f(x,r')}{g(x,r)}dt} = h(x,r) \end{aligned}$$

■

Using Claim (5), we have::

Proposition 3 *Let $F(x,r)$ represent the class of distributions with common support $[a,b]$, indexed by r as above ($r' > r \Rightarrow F(x,r') \succ_{MLRC} F(x,r)$). If $h(x,r)$ is monotonically increasing (i.e. the class of distributions have monotone increasing hazard rates) then the optimal threshold value is increasing in r i.e. $\theta^*(r') > \theta^*(r)$ $\forall r' > r$.*

Proof. The monopolists problem in the first stage is to

$$\max_{\theta} \theta \int_{\theta}^{b} f(s,r)ds$$

where θ is the threshold value. Since the maximand approaches 0 at either limit and is positive and continuous in between, the function attains a maximum. The first order conditions for this maximum, using Leibnitz's rule is given by

$$-\theta^* f(\theta^*,r) + 1 - F(\theta^*,r) = 0 \quad (14)$$

$$\text{or, } \theta^*(r)h(\theta^*,r) = 1 \quad (15)$$

Differentiating this function implicitly we find that

$$\frac{\partial \theta^*}{\partial r} = -\{\frac{\theta^* h_r(\theta^*,r)}{h(\theta^*,r) + \theta^* h_\theta(\theta^*,r)}\}$$

The monotonically increasing hazard rate implies $h_\theta(\theta^*,r) > 0$. Thus the overall sign of the expression is the opposite sign of $h_r(\theta^*,r)$. Together with the previous claim, this proves the desired result: As r increases, the hazard rate decreases and therefore the threshold value for the optimal contract increases. ■

4.3 Discussion: Changes in Heterogeneity and Incomplete Information

To discuss this result it is useful to focus on the marginal conditions implied by (13). By slightly decreasing her threshold, the monopolist's gain is the product of the density of consumers at the threshold and the value of the threshold itself, $\theta f(\theta)$. The loss is the revenue from all individuals who were earlier receiving the good and paying a higher price. Since our marginal willingness to pay function, $p(x;v)$, explicitly imposes that the change in price from including a marginal type is exactly 1, this loss is simply $(1 - F(\theta))$. Combining the marginal conditions leads to the first order condition given by (15).

This link between the hazard rate and the threshold suggests that results *independent* of the original value of the threshold can only be obtained for changes that lead to increases (or decreases) in the hazard rate at all points of the distribution. This requirement severely restricts the class of distribution changes that can be studied since $h(x,r') \leq h(x,r)$ at *all* x implies that $F(x,r')$ first order stochastically dominates $F(x,r)$ and hence that $E(X,r') \geq E(X,r)$.[15] Hence, *all* distributional changes that satisfy this requirement necessarily imply a change in the mean of the underlying distribution, ruling out for instance, mean preserving spreads in the willingness to pay.

An increases in the mean willingness-to-pay unambiguously increases the value of the threshold only if the change satisfies the MLR condition. This does not, however, translate into an automatic decrease in the mass of consumers receiving the good—whether or not this happens depends on the shape of the hazard rate function at the threshold. We wish to sign $\frac{\partial}{\partial r}\int_a^{\theta^*(r)} h(t,r)dt$ since (see footnote 15) the mass of individuals receiving the good, $1-F(\theta,r) = \frac{1}{\exp(\int_a^{\theta^*(r)} h(t,r)dt)}$. An increase in $\int_a^{\theta^*(r)} h(t,r)dt$ thus implies that there is a *decline* in the mass receiving the good. Implicitly differentiating:

$$\begin{aligned} & \frac{\partial}{\partial r}\int_0^{\theta^*(r)} h(t,r)dt \\ = & \frac{\partial \theta^*}{\partial r}h(\theta,r) + \int_0^{\theta^*(r)} \frac{\partial h(t,r)}{\partial r}dt \end{aligned} \tag{16}$$

Following an MLR change (as indexed by an increase in r), the first part of this expression that represents a "price" effect is unambiguously positive while the second, which represents a "quantity" effect, is negative. If the sum of these terms is positive, the threshold increase would be associated with greater exclusion. However, since the shape of $h(x,r)$ will in itself affect $\theta^*(r)$, exercises based on *ceteris paribus* assumptions yield no further insight. This is further clarified through the graphical exposition in Figure 2.

This two part graph shows the simultaneous determination of the threshold and (a function of) the mass of individuals receiving the good. The bottom half of the graph shows the hazard rate functions of two distribution functions, $h(x,r)$ and $h(x,r')$ (note that the area under these functions is proportional to the mass of individuals *excluded*). The top half shows the determination of the threshold in the following sense: Given the first order condition, $\theta^* h(\theta^*,r) = 1$, the threshold is exactly the fixed point of the inverse mapping $\theta \to \frac{1}{h(\theta)}$ i.e. its determined by the intersection of the inverse hazard function, $\frac{1}{h(x,r)}$ and the 45^0 line. In the graph above, the thresholds associated with $h(x,r)$ and $h(x,r')$ are thus exactly T and T'.

The comparison that we are interested in is the area under the hazard rate function, evaluated between 0 and the threshold. For $h(x,r)$ this is given by the shaded region (DATO) and for $h(x,r')$ by the region (DBT'O). The trade-off is then represented by the region shaded in black and that shaded in light gray and note that, for infinitesimal changes the region in light gray is exactly the first term in the expression above, $\frac{\partial \theta^*}{\partial r}h(\theta,r)$,and that in black is the (negative) second term, $\int_0^{\theta^*} \frac{\partial h(t,r)}{\partial r}dt$. Consequently, individuals will be excluded from receiving the good if the increase in area due to a higher 'price' (threshold) is greater than

[15] Since $\frac{f(x,r')}{1-F(x,r')} \leq \frac{f(x,r)}{1-F(x,r)} \Rightarrow \int_a^x \frac{f(x,r')}{1-F(x,r')} \leq \int_a^x \frac{g(x,r)}{1-G(x,r)}$
$\Rightarrow -\ln(1-F(x,r')) \leq -\ln(1-F(x,r))$
$\Rightarrow F(x,r') \leq F(x,r)\forall x$. Note also that from the second expression, we immediately obtain that $(1-F(x,r)) = e^{-\int_a^x h(t,r)dt}$ since $F(a)=0$ by definition of the support.

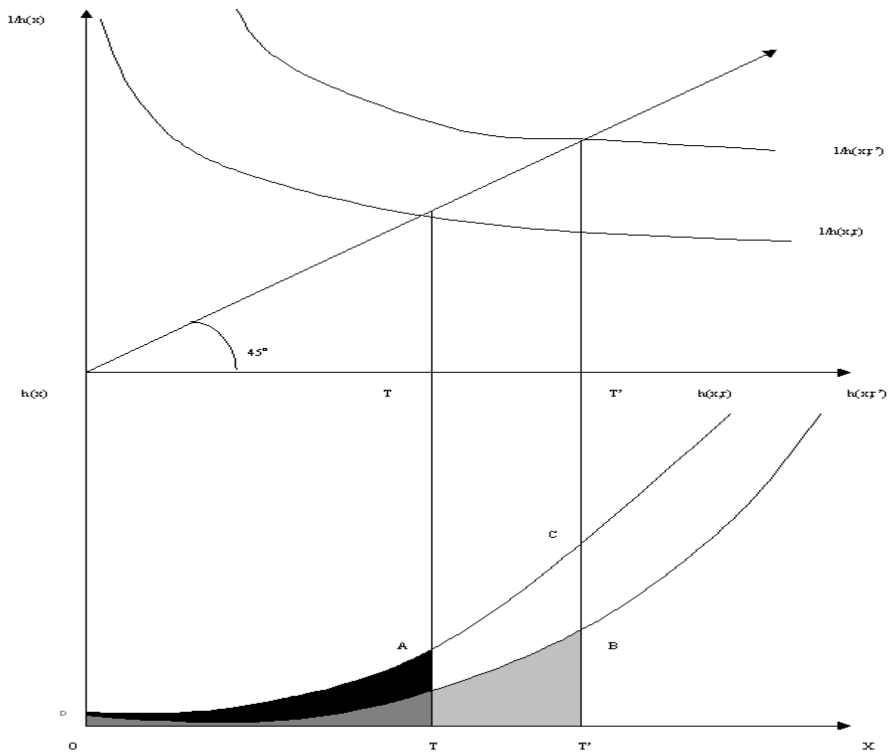


Figure 2: Hazard Rate, Thresholds and Mass Receiving the Good

the decline due to a lower 'quantity'—which effect dominates depends critically on the slope of the hazard rate function at the point of the threshold: It is easy to see for instance, that in the case of the uniform, increases in the threshold (decreases in the hazard rate) leads to *less* exclusion, while for the exponential the mass of recipients remains unchanged with changes in the threshold.

Some more intuition can be provided for when we expect decreases in the hazard rate to lead to greater exclusion. In particular, examining the two terms (areas) again, we see that exclusion is more likely when the hazard rate is relatively large but changes in the hazard rate (with respect to r) are relatively small. Further, note that the hazard rate at x is simply the ratio of the mass at x, $f(x)$, to the mass on the right of $x, (1 - F(x)$. Thus for distributions of types which are relatively skewed this exclusion condition is likely to be met at high threshold values. In other words, it is likely that in such unequal environments where there is already a high level of exclusion, further MLRC increases, despite raising the mean willingness to pay, may in fact result in even further exclusion of people (i.e. both the threshold value and the mass of people excluded would increase).

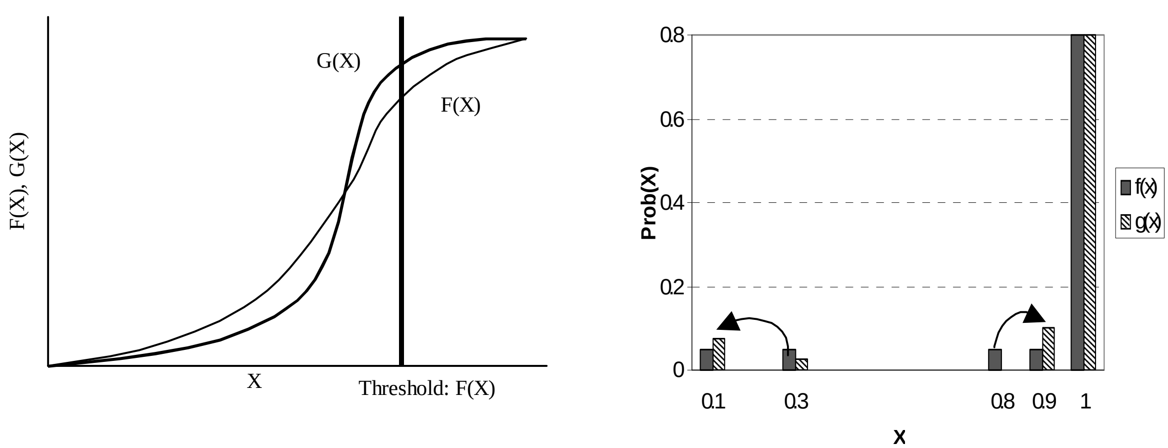
This link between the threshold and the hazard rate implies that more general results are difficult to obtain: Any result that is *independent* of the initial threshold must necessarily depend on an increase (or decrease) in the hazard rate *all points* of the support. Consequently, unambiguous results with regard to the threshold can only arise when we restrict ourselves to distributional changes that result in a decrease (or increase) in the hazard rate throughout the support of the distribution.

In general though, changes in the threshold and the mass of individuals receiving the good will depend on changes in the hazard rate in the *locality* of the threshold. To illustrate this further, it is useful to construct an example that demonstrates the sensitivity of θ^* and $1 - F(\theta^*)$ to the initial location of the threshold. We consider another class of distributional changes that are widely studied due to their direct relationship with inequality (through Lorenz dominance). Specifically, we show that mean-preserving spreads can have ambiguous effects on *both* the threshold and the mass of recipients depending on the initial level of the threshold.[16]

4.3.1 Example: Mean Preserving Spreads Yield Ambiguous Results?

The absence of general results regarding changes in the provision of the public good with a mean-preserving spread (MPS) arises from very limited restrictions that such spreads place on changes in the distribution. Specifically, an MPS requires only that mass is shifted out of the more 'central' parts of the distribution and moved more towards the 'tails'—there is however, no restriction on the exact manner in which such a transfer could occur. It is fairly easy to construct cases where the threshold increases and the mass of individuals receiving the good decreases following a mean preserving spread—in fact, for both the uniform and the normal, both these conditions are satisfied generically.[17] However, it is also possible that a mean preserving spread has exactly the opposite impact—this case arises if there is a significant shift in the mass of the distribution to a point just left of the initial threshold. As the graph on the left in figure 3 demonstrates, in this case the monopolist may decrease the optimal threshold to take advantage of this increased mass.

To make this intuition more concrete, consider the following discrete distributions, $f(x)$ and $g(x)$ on [0,1], depicted in Figure 3:



Mean Preserving Spread Leads to Decrease in Threshold

X	1/10	3/10	8/10	9/10	1
f(x)	2/40	2/40	2/40	2/40	32/40
g(x)	3/40	1/40	0	4/40	32/40

[16] Note that a mean-preserving spread immediately rules out a decrease (or increase) in the hazard rate at all points of the support since, as noted above, the only distributional changes with an unambiguous increase (decrease) in the hazard rate must satisfy (strict) first-order stochastic dominance.

[17] Proofs and simulations available with the authors.

From $f(x)$ it is possible to construct a single MPS which gives rise to the distribution $g(x)$ on the right above. The distributions $g(x)$ and $f(x)$ satisfy the conditions for differing by a single mean preserving spread: there is a change in the mass of the distribution at four points, say a_1, a_2, a_3 and a_4 s.t.

1. $g(a_1) - f(a_1) = -\{g(a_2) - f(a_2)\}$ and $g(a_3) - f(a_3) = -\{g(a_4) - f(a_4)\}$ and,
2. the mean of the distribution remains unchanged (Rothschild and Stiglitz 1970).

For the distribution $f(x)$, the monopolist maximizes her profit by charging the price '1', whereas for the distribution $g(x)$, profits are maximized by charging the lower price '9/10'. Hence, following a mean preserving spread, there is a decline in the posted price of the commodity, and an increase in the mass (from 32/40 to 36/40) of individuals receiving the good. The essence of this counter example is that an MPS need not shift mass to the absolute tails of the distribution, but only away from the center. Thus, if the monopolist was initially to the right of a point which sees a sudden increase in mass following the MPS, she may well decrease her threshold and increase the mass of individuals receiving the good i.e. greater heterogeneity leads to an *increase* in the provision of the commodity.

5 Conclusion

We have considered the problem of a monopolist supplying an excludable, non-rival good to a population characterized by consumers who have differing willingness to pay for the commodity. Our first contribution was to show that when consumer types are unobservable, under certain restrictions on preferences, a "threshold" result is obtained that simplifies the characterization of the optimal contract. Under these restrictions, the optimal contract offered by the monopolist is characterized by a threshold—those above the threshold receive the same positive amount of the good, while those below are not provided with the commodity.

The second contribution relates the provision of the public good to changes in the distribution of willingness to pay in the population. Under perfect information, the amount of the good provided depends only on the mean of the distribution, with all consumers receiving identical quantities. However, when consumer types are unobservable, this neutrality result breaks down. In an economy with two consumer types, increases in inequality unambiguously lead to a higher threshold price and greater exclusion for those with a low willingness to pay. In economies with multiple consumer types, general results are not readily obtained and depend on specific distributional changes. In particular, for distributions with increasing hazard rates, any distributional change that satisfies the monotone likelihood ratio condition necessarily leads to an increase in the threshold that the monopolist sets for the good. The degree of exclusion may increase or decrease, depending on the shape of the hazard rate function at the threshold.

We also present the structure of the monopolist's decision problem in the context of the hazard rate of the distribution. The structure is readily amenable to empirical analysis where the distribution and the distributional change can be clearly specified—one of the main contributions of this paper has been to show that the extension of the monopolist's problem to the imperfect information problem allows for a clear separation into the threshold decision followed by the quantity decision. Examining the impact of a specific distributional change is then identical to examining a change in the willingness to pay distribution for a *standard* monopolist facing unit demand from each agent type. This simplification of the problem allows for considerable gains in applied research on environments with monopolistic provision under imperfect information. Recent examples include the provision of commodities over the internet available as electronic downloads such as music, movies, journal articles and net-courses.

This discussion also helps explain where threshold contracts are likely to arise. While a number of internet download packages are indeed characterized by threshold contracts specifying a single price and number of permissible downloads, there are also examples of firms offering multiple price-quantity contracts. One argument for why this is the case is that the ratio-monotonicity condition is violated so that some individuals have high willingness to pay for small quantities of the commodity, while others have a higher willingness to pay only for additional units.

References

[1] Athey, S. 2002. Monotone comparative statics under uncertainty. Quarterly Journal of Economics 117(1), 187-22.

[2] Bardhan, P., Ghatak, M., Karaivanov, A. 2002. Inequality, market imperfections, and the voluntary provision of collective goods. Mimeo.

[3] Baron, D., Myerson, R.B. 1982. Regulating a monopolist with unknown costs. Econometrica 50, 911-930.

[4] Bergstrom, T., Blume, L., Varian, H. 1986. On the private provision of public goods. Journal of Public Economics 29, 25-49.

[5] Bernheim, B.D. 1986. On the voluntary and involuntary provision of public goods. American Economic Review 76(4), 789-93.

[6] Brito, D.L., Oakland, W.H. 1980. On the monopolistic provision of excludable public goods. American Economic Review 70(4), 691-704.

[7] Cornes, R., Sandler, T. 1994. The comparative static properties of the impure public good model. Journal of Public Economics 54, 403-21.

[8] Hellwig, M.F. 2003. Public-good provision with many participants. Review of Economic Studies 70 (3), 589-614

[9] Itaya, J., de Meza, D., Myles, G.D. 1997. In praise of heterogeneity: public good provison and income distribution. Economic Letters 57, 289-296.

[10] Laffont, J.J., Tirole, J. 1993. A Theory of Incentives in Regulation and Procurement. The MIT Press, Cambridge.

[11] Maskin, E., Riley, J. 1984. Monopoly with incomplete information. Rand Journal of Economics 15(2), 171-195.

[12] Milgrom, P.R. 1981. Good news and bad news: representation theorems and applications. Bell Journal of Economics 12, 380-391.

[13] Norman, P. 2003. Efficient mechanisms for public goods with use exclusion. Review of Economic Studies, forthcoming.

[14] Rothschild, M., Stiglitz, J.E. 1970. Increasing risk I: a definition. Journal of Economic Theory 2, 225-243.

[15] Rothschild, M., Stiglitz, J.E. 1971. Increasing risk II: it's economic consequences. Journal of Economic Theory 3, 66-84.

[16] Rothschild, M., Stiglitz, J.E. 1973. Some further results on the measurement of heterogeneity. Journal of Economic Theory 6, 188-204.

[17] Sen, A., Foster, J.E. 1997. On Economic Heterogeneity. Clarendon Press, Oxford.

[18] Schmitz, P.W. 1997. Monopolistic provision of excludable public goods under private information. Public Finance/Finances Publiques 52(1), 89-101.

[19] Warr, P.G. 1983. The private provision of a public good is independent of the distribution of income. Economic Letters 13, 207-11.

Appendix: Proofs

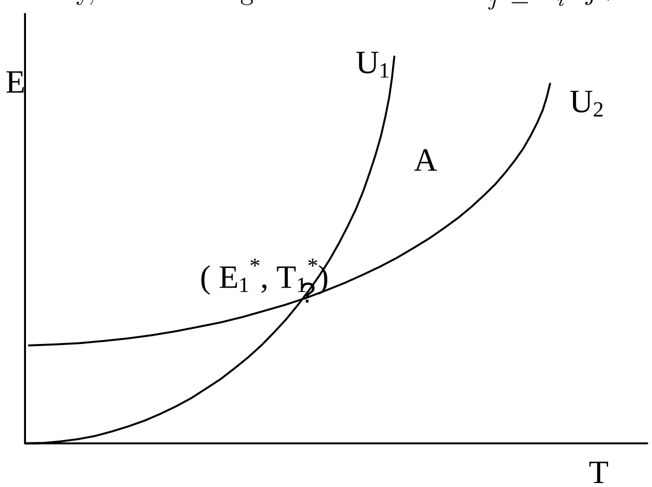
Claim 2: *If a consumer of a given type weakly prefers a contract with higher public good to another, any consumer of a higher type will strictly prefer the former to the latter contract.*

$$\forall E_2 \geq E_1 \& \quad \langle E_1, T_1 \rangle \neq \langle E_2, T_2 \rangle, \quad \langle E_1, T_1 \rangle \precsim_i \langle E_2, T_2 \rangle \Rightarrow \langle E_1, T_1 \rangle \prec_j \langle E_2, T_2 \rangle \ \forall j > i$$

Proof. This claims follows from the fact that for any given public good level, a consumer of higher type is always willing to pay more than a consumer of lower type. More specifically, the slopes of the consumer's indifference curves can always be type-ordered and the ordering remains that same regardless of the public good level. In the above setup, we get:$\frac{dT}{dE}|_{dU_i=0} = v_i g'(.)$. This implies that, for any given level of public good received $\frac{dT}{dE}|_{dU_j=0} > \frac{dT}{dE}|_{dU_i=0} \forall j > i$. Therefore it must be that if a consumer of type i weakly prefers contract 2 with higher E to contract 1, then the indifference curve of any higher type j through contract 1 must lie below contract 2 and hence type j strictly prefers contract 2 to contract 1. ■

Claim 3: *Any optimal menu of contracts offered by the monopolist, $\langle E_i^*, T_i^* \rangle_{i \in \Lambda}$, has to be such that the contract chosen by the higher consumer type offers an equal or higher level of public good. i.e. $E_j^* \geq E_i^* \forall j > i$, where $\langle E_i^*, T_i^* \rangle$ denotes i's preferred contract choice from the menu of contracts offered.*

Proof. Let $\langle E_1^*, T_1^* \rangle$ be consumer type 1's optimal contract choice. This implies that $\langle E_1^*, T_1^* \rangle \succsim_1 \langle E_i^*, T_i^* \rangle \forall i$ i.e. all other contracts offered in the menu must lie either on or below Type 1's indifference curve (IC) through $\langle E_1^*, T_1^* \rangle$. Now consider the Type 2 consumer. From above we know that Type 2's IC through point $\langle E_1^*, T_1^* \rangle$ will be as drawn in Figure 4 below i.e. the slope of the IC is lower for type 2 at point $\langle E_1^*, T_1^* \rangle$. Now consider Type 2's optimal contract choice $\langle E_2^*, T_2^* \rangle$. Since this is his optimal choice it must be that $\langle E_2^*, T_2^* \rangle \succsim_2 \langle E_1^*, T_1^* \rangle \ \forall i$ i.e. $\langle E_2^*, T_2^* \rangle$ must lie on the same or higher IC for Type 2. Together these two conditions imply that $\langle E_2^*, T_2^* \rangle$ must lie in the shaded area shown in the figure and hence $E_2^* \geq E_1^*$. More generally, the same argument shows that $E_j^* \geq E_i^* \forall j > i$



■

Higher Types Receive Higher E

Claim 4: *If $\langle E_i^*, T_i^* \rangle_{i \in \Lambda}$ represents the optimal (for the monopolist) implementable menu of contracts, then it must be that for any given consumer type, his downward incentive constraint will always be binding i.e. $\langle E_i^*, T_i^* \rangle \sim_i \left\langle E_{i-1}^*, T_{i-1}^* \right\rangle \forall i$*

Proof. Suppose that the above statement is not true i.e. $\langle E_i^*, T_i^* \rangle \succ_i \left\langle E_{i-1}^*, T_{i-1}^* \right\rangle$. Since $\left\langle E_{i-1}^*, T_{i-1}^* \right\rangle$ is by definition the optimal contract for type $i-1$, it must be that $\left\langle E_{i-1}^*, T_{i-1}^* \right\rangle \succsim_{i-1} \langle E_k^*, T_k^* \rangle \ \forall k$ and in particular for $k = 1, .., i-1$. Using the above two claims this then implies that $\left\langle E_{i-1}^*, T_{i-1}^* \right\rangle \succ_i \langle E_k^*, T_k^* \rangle$ $\forall k = 1, .., i-1$. Now consider an alternate menu of contracts such that:

$$\widehat{E}_i = E_i^*$$
$$\widehat{T}_i = \begin{cases} T_i^* & for \ k < i \\ T_i^* + \delta & for \ k \geq i \end{cases}$$

where $\delta > 0$. Note that given our assumption that $\langle E_i^*, T_i^* \rangle \succ_i \left\langle E_{i-1}^*, T_{i-1}^* \right\rangle$, we can always choose a small enough δ s.t. $\left\langle \widehat{E}_i, \widehat{T}_i \right\rangle \succ_i \left\langle \widehat{E}_{i-1}, \widehat{T}_{i-1} \right\rangle$ and hence we know that $\left\langle \widehat{E}_{i-1}, \widehat{T}_{i-1} \right\rangle \succ_i \left\langle \widehat{E}_k, \widehat{T}_k \right\rangle \forall k = 1, .., i-1$. Thus under the new menu of contracts offered, consumer Type i continues to prefer contract $\left\langle \widehat{E}_i, \widehat{T}_i \right\rangle$ i.e. he does not want to move to a contract with a lower E value despite having to pay more under this new menu of contracts. Similarly Type i will not want to move to any contract with a higher E value as that would violate $\langle E_i^*, T_i^* \rangle$ having been optimal (in comparison to higher types there has been no effective change under the new menu of contracts as the incentive constraints remain the same i.e. $-\delta$ is added to both sides of any incentive constraint between type I and a higher type). Thus i continues to choose $\left\langle \widehat{E}_i, \widehat{T}_i \right\rangle$. Similarly, we can show that all other types l continue to prefer contract $\left\langle \widehat{E}_l, \widehat{T}_l \right\rangle$. For any type $j > i$, using the above two claims and that $\left\langle \widehat{E}_i, \widehat{T}_i \right\rangle \succ_i \left\langle \widehat{E}_k, \widehat{T}_k \right\rangle \forall k < i$ we know that $\left\langle \widehat{E}_i, \widehat{T}_i \right\rangle \succ_j \left\langle \widehat{E}_k, \widehat{T}_k \right\rangle \forall k < i \textit{ and } \forall j > i$. Thus type $j > i$ will not want to switch to any contract with a lower E. For contracts with higher E, since the relevant ICs have not changed they will not want to switch to such contracts either. Finally no type $k < i$ will want to switch. For a contracts with higher E, they now have to pay even greater fees and so if they didn't prefer such contracts under the old menu, they will definitely not prefer them now. For contracts with lower E, the new menu of contracts is exactly the same and so type $k < i$ will optimally choose $\left\langle \widehat{E}_k, \widehat{T}_k \right\rangle$. Thus we have shown that under the new menu of contracts each type i chooses contract $\left\langle \widehat{E}_i, \widehat{T}_i \right\rangle$ and therefore there is no change in the amount of public good provided in comparison will the old menu of contracts. However it is clear that the monopolist's total fees collected have increased which contradicts our initial assumption that $\langle E_i^*, T_i^* \rangle_{i \in \Lambda}$ was optimal for the monopolist. Therefore it must be that $\langle E_i^*, T_i^* \rangle \sim_i \left\langle E_{i-1}^*, T_{i-1}^* \right\rangle \forall i$. ■