

Public Sector Rationing and Private Sector Selection

Simona Grassi

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Madrid)
España

sgrassi@eco.uc3m.es

Ching-to Albert Ma

Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
USA

ma@bu.edu

Incomplete, preliminary. Do not quote, cite, or circulate.

June 2008

Acknowledgement: We thank seminar participants at Boston University, Carlos III University in Madrid, Harvard University, the Health Economic Policy conference in Crans Montana, Switzerland, 9th European Health Economics Workshop in Bergen for their comments and suggestions. The first author thanks the Italian Fulbright Foundation for financial support.

Abstract

We study the interaction between a public sector and a private sector in the provision of a private good. The public sector has a limited budget. A private firm may supply the good to those consumers who are rationed by the public system. Consumers have different amounts of wealth, and costs of providing the good to them vary. We consider two information regimes: first, the public supplier observes only wealth information; second, the public supplier observes both wealth and cost information. The public supplier chooses a rationing policy based on its information; simultaneously, the private firm, observing only cost but not wealth information, chooses a pricing policy. In the first information regime, there is a continuum of equilibria; in each, rich consumers are rationed, and the private firm sells to these rationed consumers at high prices. In the second regime, there is a unique equilibrium. The public supplier allocates the good to consumers according to a cost-effectiveness rule. Rationed consumers have high costs relative to the benefit. The private firm sells to wealthy rationed consumers at high prices.

1 Introduction

Many governments and public organizations provide or subsidize goods and services such as education and health care. Public provision at subsidized or zero charge to consumers often coexists with a private market. In this paper, we model the interaction between the public and private sectors in the provision of an indivisible good. In our model, the private sector reacts to public supplies, and simultaneously the public sector reacts to prices in the private market. We derive equilibria of the game between the two sectors under various information regimes.

Our model is set up to address a number of issues. First, the private sector may react to public supply by selecting or cream skimming profitable consumers. How does the public sector react to cream skimming? Second, a consumer enjoys different surpluses from public and private supplies. How does this affect the equilibrium allocation from the public sector and prices in the private sector? Third, the allocation in the public sector and the supply in the private sector may be based on different information. How do equilibria change when information structures change?

Consumers in our model differ in two dimensions. They have different wealth levels. The costs of providing the good to them also differ. We use the wealth heterogeneity to model consumers' differences in valuations of the good. Rich consumers are more willing to pay for the good than poor consumers. The cost heterogeneity dimension arises because consumer characteristics may determine how much it costs to supply the good to them. For example, in the education market, the cost of helping a student to achieve a given academic standard depends on the student's ability and aptitude. In the health care market, the cost of a treatment depends on illness severity. Variations in these characteristics affect the costs of provision.

We use a monopoly model for the private sector, but the analysis for a Cournot private sector remains the same; the analysis also extends easily to a perfectly competitive private sector. The firm in the private sector observes a consumer's cost but not his wealth. Because consumers have different willingness to pay, the firm faces a downward-sloping demand function. By lowering the price, the firm sells to consumers with lower wealth and lower willingness to pay. As a consumer's cost changes, the firm will change its profit-maximizing price.

Public supply is assumed to be free of charge to consumers but the available budget is insufficient to cover all consumers. The public supplier must use a rationing rule. We consider two information regimes. In the first, consumers' wealth information is available to the public supplier, and the rationing rule is based on it. In the second, consumers' wealth and cost information is available, and the rationing rule is based on both pieces of information.

The game between the public supplier and the private firm is described as follows. First, consumers' wealth and cost variables are realized as draws from their independent distributions. In each information regime, the public supplier sets a rationing rule. Simultaneously, the private firm sets a price as a function of consumers' costs. Those consumers who are rationed by the public supplier may choose to buy the good from the private firm at the offered price. The public supplier maximizes the sum of consumer utility given the budget constraint and the pricing rule in the private sector. The private firm maximizes profit given the public supplier's rationing rule.

Our game is used to study the three issues raised above. Cream-skimming and selection are captured by the private firm's pricing strategy. Prices are generally above marginal costs, and are responses against the public provider's rationing policy. Consumers who obtain the good from the public supplier are no longer potential customers for the private firm. For example, if the public supplier serves all consumers with wealth below a threshold, the private firm now realizes that there are no poor consumers in the market. Even as costs drop to very low levels, the private firm has no incentive to reduce prices, because all available consumers have high willingness to pay.

The private firm's price schedule depends on consumers' costs. Among consumers who are willing to purchase, wealthy ones obtain higher inframarginal surplus. Because the public supplier attempts to maximize total consumer utility, it has an incentive to supply poor consumers and ration wealthy consumers. Letting wealthy consumers buy from the private market generates more surplus. This trade-surplus effect captures the public supplier's basic incentives when reacting against the private sector's pricing strategy.

In the first information regime, when rationing is based on wealth information, the private firm's pricing strategy and the trade-surplus effect reinforce each other. There is a continuum of equilibria; in each, the

public supplier always rations wealthy consumers. The private firm sets a price at least as high as one it would set if the public supplier did not have any budget. In the equilibrium that generates the highest consumer utility, consumers obtain the good from the public supplier if and only if their wealth is below a threshold, and the private firm's price will remain high even when the consumer's cost is low. Wealthy consumers who have low costs will be hurt by the public supply.

In the second information regime, rationing is based on both wealth and cost information. Here, a cost effectiveness principle applies: the public supplier allocates the good to consumers when the benefit is above the cost (adjusted by the shadow price of the budget). In our model, absent the private sector, the optimal rationing scheme supplies the good to consumers whose costs are below a threshold (and the rationing rule is independent of wealth). The cost effectiveness principle continues to apply when the private market is reactive. In the unique equilibrium, consumers are allocated the good by the public supplier if and only if their costs are below the same threshold. The private firm only sells to the remaining consumers, at prices that it would have charged if there was no public supply. It is as if the public supplier has ignored the private firm. While low-cost consumers are served by the public sector irrespective of their wealth, only rationed and wealthy consumers buy from the private firm. Low-cost consumers are better off as a result of the public supply because they get the good for free; high-cost consumers are not hurt by the public supply.

Each of the three players of our model—consumers, the private firm, and the public supplier—plays an important role. The consumers' wealth and cost heterogeneities imply differences in valuations of the good, and these differences generate a downward sloping demand function. We use a separable utility function for the analysis, and discuss in a later section more general utility functions. We also assume that valuations, and hence demands, are independent of costs. This assumption makes the analysis tractable, but we also discuss how it may be relaxed.

We consider two information regimes for the extensive form. The private firm always observes costs, but we let the public supplier observe wealth in one regime, and wealth and cost in the other. The assumption that a firm knows its cost seems natural. We also think that the government should have good information about consumers' wealth or income through tax returns. This is the first information regime. Cost information

also may be available to the public supplier. Afterall, the public supplier actually may provide the good to the consumers. This is the second information regime.

We have included the first information regime because often a public supplier may not be able to use cost information. In the health and education markets, the government policy may commit itself to supply services at all reasonable costs. Furthermore, there may be some decentralization in the implementation of provision. In the health care market, for example, clinicians may decide on medical services based on needs rather than costs. In the education market, school districts are committed to provide education to all eligible students irrespective of their resource requirements. In both cases, the rationing issue concerns who are entitled to public provision. In our model, this implies that rationing policy must be based only on wealth. That is, the public supplier grants the right to service to consumers according to their wealth levels, and assumes the costs.

The private firm reacts against the public supplier's rationing policy by setting a price as a function of cost. A rationing policy affects the portfolio of consumers that are in the market. When setting a profit-maximizing price, the firm must take this into account. We have chosen to use a monopoly model for easy exposition; a Cournot model will yield the same results. The important structure of the private sector for our analysis is that it is responding strategically against the rationing policy. By contrast, in a perfectly competitive private sector, prices equal to marginal costs, irrespective of the rationing policy; this is inadequate for our purpose.

The public supplier's policy instrument is a rationing rule. Nonprice rationing is ubiquitous. Many governments set prices of education and health care to zero, and ration these services in various ways. Presumably, this may be due to fairness or political considerations. There are also economic reasons for favoring rationing. For the health market, insuring consumers' financial risks due to illness is a fundamental goal of a benevolent government. Under social insurance, consumers should not be exposed to too much financial risks upon becoming sick. Setting negligible prices of health care is a form of risk sharing. Some rationing policy must then be used. For the education market, a government may encourage the investment of human capital, which may enhance economic growth and create externalities. Again, reducing costs of

education to zero and then adopting some rationing policy may be a sensible policy.

Eliminating monetary subsidies as the public supplier's policy instrument, we abstract from redistribution issues. We have in mind a hierarchical system where redistribution of income is handled by general or income taxation. The public supplier in our model is not part of the agency responsible for income redistribution. It uses a given budget to provide the good to some consumers, and is unable to make further changes in income distribution by setting (subsidized) prices.

The strategic interaction between the public supplier and the private firm is set up as a simultaneous-move game. This extensive form imposes some symmetry on commitment between the two players. Generally, the public supplier would enjoy becoming a Stackelberg leader by committing to a rationing rule. However, the public supplier would renege on the Stackelberg policy. In other words, the public supplier's Stackelberg rationing rule is time inconsistent.

The commitment issue is this. If the public supplier rations some poor consumers, the portfolio of consumers in the market consists of both poor and wealth consumers. When the firm observes a low-cost consumer, it may be profit-maximizing to sell at a low price. Prices would then fall as costs drop, rather than stay constant. This would benefit rich consumers. But given the strictly increasing price schedule, the public supplier would like to ration more consumers with higher wealth and supply more consumers with lower wealth. Rationing the poor is never a best response given a strictly increasing price schedule.

We have not taken a mechanism design approach to derive an optimal regulatory policy for the monopolist. Our concern is on the price reaction of a private sector which is imperfectly competitive. We assume a single firm for the formal analysis for ease of presentation. How to regulate an imperfectly competitive industry is arguably less well understood. Whether such an industry should be regulated is unclear either. Besides, we already limit the policy instrument available to the public supplier to a rationing rule. Regulatory policies likely should require more instruments. We therefore stick with an explicit game form between the public supplier and the private firm. In any case, the game form seems natural and adequate for studying strategic interaction between these sectors.

Many papers in health economics (for example Barros and Olivella [2], Ellis [8]), study selection and

crowding out. They focus on cream skimming, which is the selection of low cost (low severity) patients. Barros and Olivella focus on waiting lists and patient selection. They show that despite private physicians selecting the lowest cost cases from those on the public waiting list, very strict or very loose rationing policies may only lead to partial cream skimming. Ellis studies the effect of different reimbursement schemes on cream skimming (and dumping.) Hoel [9] considers the coverage of public health insurance when consumers may buy from a competitive market. There is no analysis about any interaction between the public and private sectors; the private sector is modelled as providing a fixed outside option for consumers.

Section 2 and its subsections lay out the components of the model. In Section 3, we describe the private firm's price responses. In Sections 4 and 5, we derive equilibrium rationing and pricing schemes for the two information regimes, respectively. Section 6 considers alternative assumptions and robustness. The last Section draws some conclusions.

2 The Model

We begin with the description of consumers in the model. Next, we introduce a private sector and then a public sector. We complete the model description by the extensive-form games between the consumers, public, and private sectors.

2.1 Consumers and willingness to pay

There is a set of consumers. Each consumer may consume at most one unit of an indivisible good. We let there be a continuum of these consumers, with total mass normalized to 1. Each consumer is indexed by two parameters, w and c . The variable w denotes the consumer's wealth. The variable c denotes the cost of supplying this good to the consumer. The cost c of provision is identical whether the good is supplied by the public or private sectors; we do not consider any productive comparative advantage between the private and public sectors in order to focus on information problems. We often use the term consumer (w, c) to refer to one who has wealth w and cost parameter c .

The variables w and c are assumed to be independently distributed. Let $F : [\underline{w}, \bar{w}] \rightarrow [0, 1]$ be the distribution function of w . We assume that F is differentiable, and let the corresponding density be f .

Similarly, let $G : [\underline{c}, \bar{c}] \rightarrow [0, 1]$ be the distribution function of c . We also assume that G is differentiable, and let the corresponding density be g . The domains of these two distributions are strictly positive and bounded. We will also assume that the densities f and g are both strictly positive over their domains. This implies that the distribution functions F and G are both strictly increasing.

For a general specification of preferences, we can let a consumer's utility be $U(w, 0)$ when he does not consume the good, and $U(w - p, 1)$ when he consumes the good at a price $p \geq 0$. The utility function U is strictly increasing, and strictly concave in w , and $U(w, 1) > U(w, 0)$. It saves on notation and simplifies the analysis if we let the utility function U be separable in the two arguments. The separability assumption says that a unit of the good generates the same utility increment, independent of the consumer's wealth. Hence, if the consumer with wealth w pays a price p to consume the good, his utility is $U(w - p) + 1$. The consumer's utility is $U(w)$ if he does not consume the good. We consider a more general specification of the utility function in Section 6.

Consider a consumer with wealth w . If this consumer is indifferent between paying a price τ to consume the good and the status quo, we have:

$$U(w - \tau) + 1 = U(w). \tag{1}$$

This equation implicitly defines a willingness-to-pay function $\tau : [\underline{w}, \bar{w}] \rightarrow \mathbb{R}_+$ for consumers with various wealth levels. Since U is concave and hence almost everywhere differentiable, the willingness to pay function is differentiable. From total differentiation of (1), we have

$$\frac{d\tau}{dw} = 1 - \frac{U'(w)}{U'(w - \tau)} > 0. \tag{2}$$

A consumer's willingness to pay for the good is strictly increasing in wealth due to the strict concavity of U .

We can illustrate our description of consumer preferences and costs with examples in the health market. The good may refer to a surgical procedure (for example, a hip replacement). Patients differ in their illness severity levels (some hip replacements are more difficult than others). For a fixed amount of improvement in health, interpreted as a unit increment of utility (for example, the ability to walk about without pain), sicker patients require more resources, and wealthy patients are more willing to pay.

In our setup consumer preferences do not directly depend on the provision cost c . In the health care example, this means that patients with different severity levels obtain the same incremental utility from the good. One interpretation is that the good provides a standardized unit of improvement in well-being. In other situations, consumers obtain different incremental utilities depending on their severity levels. In this case, consumer preferences depend on cost, and we will discuss this alternative assumption in Section 6.

2.2 The private sector and profit-maximizing prices

We now describe a private sector. We let there be a single firm monopoly in the private sector; the cases of a Cournot private sector as well as a perfectly competitive private sector also will be considered. The selection issue we focus on is this. Facing any consumer, the private firm may observe the cost of providing a unit of the good to the consumer, but the private firm does not get to observe a consumer's wealth level. For consumer (w, c) , the private firm gets to observe only c , but not w .

The monopolist aims to maximize profit by setting a price to sell to consumers, given the cost c . For now, assume that the monopolist may have access to the entire mass of consumers. Not knowing the consumer's wealth, the monopolist does not know the willingness to pay. By setting a price (that is above cost c), the monopolist may sell to those consumers with willingness to pay higher than the price. Obviously, the monopolist will not set a price outside the range of willingness to pay τ . Setting a price p is equivalent to selecting the wealth level of the *marginal* consumer w , where $p = \tau(w)$. By the strictly monotonicity of τ , consumers with $w' > w$ has $\tau(w') > \tau(w)$, and hence are willing to pay p to purchase the good. The function τ is like a demand function; we simply restate the common principle that a monopolist may choose equivalently between a price and a quantity while respecting the demand function.

Suppose that the monopolist knows the cost c of providing the good to a consumer. If it intends to sell to consumers with wealth w or higher, it sets a price $\tau(w)$, and its profit is

$$\pi(w; c) \equiv \int_w^{\bar{w}} f(x) dx [\tau(w) - c] = [1 - F(w)] [\tau(w) - c]. \quad (3)$$

Let the profit-maximizing choice of the marginal consumer be $\hat{w}^m : [c, \bar{c}] \rightarrow [\underline{w}, \bar{w}]$,

$$\hat{w}^m(c) \equiv \operatorname{argmax}_w [1 - F(w)] [\tau(w) - c]. \quad (4)$$

The marginal consumer, one who pays the price equal to his willingness to pay, has wealth $\widehat{w}^m(c)$, and the profit-maximizing quantity is $1 - F(\widehat{w}^m(c))$. Although $\widehat{w}^m(c)$ denotes the marginal consumer, we also call $\widehat{w}^m(c)$ a quantity function when there is no possibility of confusion.

We assume that $\widehat{w}^m(c)$ is single-valued. By the Maximum Theorem $\widehat{w}^m(c)$ is continuous. We further assume that as c varies over $[\underline{c}, \bar{c}]$, the marginal consumers vary over a proper subset of $[\underline{w}, \bar{w}]$ with $\underline{w} < \widehat{w}(\underline{c}) < \widehat{w}(\bar{c}) < \bar{w}$. This requires that the variation in wealth is sufficiently large relative to the variation in costs.

Lemma 1 *The quantity function $\widehat{w}^m(c)$ is strictly increasing. If the monopolist has access to all consumers, the monopolist raises its price and sells to less consumers as cost increases.*

Proof of Lemma 1: The proof of Lemma 2 below shows that the quantity function $\widehat{w}^m(c)$ must be increasing; simply set $\theta(w)$ to 1 for all w in Lemma 2. Here we show that it is strictly increasing. The profit function (3) is differentiable in w . The first-order derivative of $\pi(w, c)$ with respect to w is $-f(w)[\tau(w) - c] + \tau'(w)[1 - F(w)]$, and this vanishes at $w = \widehat{w}^m(c)$ since $f(w) > 0$ by assumption. Hence, for $\epsilon > 0$, $-f(w)[\tau(w) - (c - \epsilon)] + \tau'(w)[1 - F(w)]$ is strictly negative at $w = \widehat{w}^m(c)$. Again because $f(w) > 0$, lowering w from $\widehat{w}^m(c)$ must strictly increase profit at cost $c - \epsilon$. In other words, $\widehat{w}^m(c)$ does not maximize $\pi(w, c - \epsilon)$. We conclude that $\widehat{w}^m(c)$ must be strictly increasing. ■

Lemma 1 does make use of the assumption $f(w) > 0$ or equivalently $F(w)$ is strictly increasing. Indeed, when all consumers are available, the private firm may always sell to more consumers by lowering its price. Hence when cost falls, it must sell to more consumers. Later we will see that if some consumers are supplied by the public sector, there may not be consumers around to accept a price reduction. Price may stop falling even when cost decreases.

2.3 The public sector and rationing

The public sector has a fixed budget B but the budget is insufficient to supply to all consumers for free. We consider two information regimes. First, only consumers' wealth information is available to the public supplier, and second, consumers' wealth and cost information is available. In each case, nonprice rationing

will be used to allocate the budget for providing the good to consumers.

Consider the first regime. The public supplier's rationing rule is defined by the function $\theta : [\underline{w}, \bar{w}] \rightarrow [0, 1]$. For any given $w \in [\underline{w}, \bar{w}]$, the public supplier provides the good to consumers with wealth below w a total of $\int_{\underline{w}}^w (1 - \theta(x))f(x) dx$ units of the good at zero cost to them, but not the remaining consumers $\int_{\underline{w}}^w \theta(x)f(x) dx$. The rationing rule θ modifies the density f so that at w , $[1 - \theta(w)]f(w)$ of consumers are supplied by the public sector at zero price, and $\theta(w)f(w)$ of consumers are available to the private market. Because the rationing rule is based on wealth, the cost c among rationed consumers is distributed according to G .

For the second regime the rationing rule is defined by the function $\phi : [\underline{w}, \bar{w}] \times [\underline{c}, \bar{c}] \rightarrow [0, 1]$. It has the same interpretation as in the first regime. For consumer (w, c) , the density $\phi(w, c)f(w)g(c)$ is available to the private firm.¹ In each regime, the supplier's objective is the sum of consumer utility, which will be defined below.

The rationing schemes θ and ϕ correspond to random rationing, but can be implemented by waiting times. We can add to the consumer preference specification a new parameter, say δ , a random variable that is independently distributed according to some distribution, say H . The utility of a consumer is now $U(w) + 1 - \delta t$ if he gets the good after a delay of t units of time. The parameter δ describes the consumer's marginal waiting cost. An impatient consumer (one with a high value of δ) may decide against the public system if he expects a long delay. By setting the delay t , the public supplier determines the fraction of consumers within a wealth group or a wealth-cost group who choose to get the good.²

2.4 Interaction between the public and private sectors

We now use the above components to set up interactions between the public and private sectors. We use a simultaneous-move game in both information regimes:

Stage 1: Nature draws (w, c) according to the densities f and g . The private firm observes c . The public

¹We restrict rationing rules to those that make the functions $\theta(w)f(w)$ and $\phi(w, c)f(w)$ integrable, so that $\int_{\underline{w}}^w \theta(x)f(x) dx$ and $\int_{\underline{w}}^w \phi(x, c)f(x) dx$ are well defined for $w \in [\underline{w}, \bar{w}]$.

²We can restrict the regulator to supply to either all or none of the consumers within a wealth class or a wealth-cost class. Rationing schemes are then functions that map $[\underline{w}, \bar{w}]$ to $\{0, 1\}$ and $[\underline{w}, \bar{w}] \times [\underline{c}, \bar{c}]$ to $\{0, 1\}$. The general rationing functions can now be interpreted as mixed strategies.

sector supplier observes either w or both w and c .

Stage 2: In each information regime, the public sector supplier chooses a rationing rule, and the private firm chooses a quantity function \hat{w} .

Stage 3: Consumers supplied by the public sector get the good for free, and consumers not supplied by the public may purchase from private firm at prices set in Stage 2.

We look for Nash equilibria in the games in the two information regimes.

2.5 Optimal rationing with an inactive private market

For now suppose that the public sector is the sole provider of the good, and the private market is inactive. Consider the first information regime where rationing can only be based on wealth. Let $\gamma \equiv \int c \, dG$ denote the expected cost. For a rationing rule θ , total consumer benefit from consuming the good is $\int_{\underline{w}}^{\bar{w}} (1 - \theta(w)) f(w) \, dw$ as each unit of consumption increases a consumer's utility by one unit. Therefore, the utilitarian welfare index is

$$V(\theta) \equiv \int_{\underline{w}}^{\bar{w}} U(w) \, dF + \int_{\underline{w}}^{\bar{w}} [1 - \theta(w)] f(w) \, dx. \quad (5)$$

The rationing rule must satisfy the budget constraint

$$\gamma \int_{\underline{w}}^{\bar{w}} (1 - \theta(w)) f(w) \, dw \leq B, \quad (6)$$

which says that the expected cost must not exceed the available budget.

The determination of a rationing rule that maximizes (5) subject to (6) is rather trivial. Any rationing rule that exhausts the budget is optimal. The public supplier allocates the good to consumers without collecting any payment. Due to the separable utility function, the utility increment is independent of w . Any rationing scheme that exhausts the budget results in the same level of the welfare index, and is optimal.

Now we consider the second information regime, where rationing can be based on wealth and cost. For a rationing rule ϕ , the welfare index is

$$V(\phi) \equiv \int_{\underline{c}}^{\bar{c}} \int_{\underline{w}}^{\bar{w}} \{U(w) + [1 - \phi(w, c)]\} f(w) g(c) \, dw \, dc. \quad (7)$$

The rationing rule must satisfy the budget constraint

$$\int_{\underline{c}}^{\bar{c}} \int_{\underline{w}}^{\bar{w}} (1 - \phi(w, c))c f(w)g(c) dw dc \leq B. \quad (8)$$

By pointwise optimization with respect to ϕ , the optimal rationing rule is given by³

$$\phi(w, c) = 0 \quad \text{for } c < c^s \quad \text{and } \phi(w, c) = 1 \quad \text{otherwise.}$$

The supplier has perfect information, and the optimal rationing rule is based on a cost-effectiveness measure. Each unit of the good yields a fixed increment of utility. The benefit relative to cost falls as cost rises. The optimal rationing policy therefore supplies the good to consumers if and only if their costs are below a threshold.⁴

3 Prices in the private market and consumer welfare

In this section we present results relating to the private firm's best response against the public sector rationing policy. We will focus on the information regime where the public sector rationing policy is based on wealth. Given a rationing rule θ , the private firm's profit from selling to consumers with wealth higher than w is

$$\pi(w; c, \theta) \equiv \int_w^{\bar{w}} \theta(x)f(x) dx [\tau(w) - c]. \quad (9)$$

which differs from the expression in (3) in that at w only a fraction $\theta(w)$ of consumers with wealth w would consider buying from the private firm. Let $\hat{w}(c)$ be the optimal quantities, and $\hat{\pi}(c)$ the maximum profit:

$$\hat{w}(c) = \arg \max_w \pi(w; c, \theta), \quad (10)$$

$$\hat{\pi}(c) = \pi(w'; c, \theta), w' \in \hat{w}(c). \quad (11)$$

For some rationing rules, there may be multiple quantities that maximize profit, so $\hat{w}(c)$ is a correspondence.

According to the Maximum Theorem, the correspondence $\hat{w}(c)$ is upper semicontinuous. An equilibrium

³The Lagrangean is $U(w) + [1 - \phi] + \lambda[B - (1 - \phi)c]$, and its first-order derivative with respect to ϕ is $-1 + \lambda c$, which is strictly positive if and only if c is larger than a threshold, say, c^s .

⁴For the general utility function, the Lagrangean is $[1 - \phi]U(w, 1) + \phi U(w, 0) + \lambda[B - (1 - \phi)c]$, and its first-order derivative with respect to ϕ is $-U(w, 1) + U(w, 0) + \lambda c$, which is strictly increasing in c , and strictly decreasing in w under the assumption $\frac{\partial U}{\partial w}(w, 1) \geq \frac{\partial U}{\partial w}(w, 0)$. Define (w^*, c^*) by $-U(w^*, 1) + U(w^*, 0) + \lambda c^* = 0$. The optimal rationing rule sets $\phi(w, c) = 0$ if and only if $w \geq w^*$ and $c \leq c^*$.

is a selection from such a correspondence. We present some monotonicity results on the firm's profit-maximization problem.

Lemma 2 *The maximum profit is strictly decreasing in c . Any selection from the profit-maximizing prices, $\hat{w}(c) = \arg \max_w \pi(w; c, \theta)$, is increasing in c ; that is, if $c_1 < c_2$, then $w_1 \leq w_2$ where $w_1 \in \hat{w}(c_1)$ and $w_2 \in \hat{w}(c_2)$.*

Proof of Lemma 2: For $c_1 < c_2$, let $w_1 \in \hat{w}(c_1)$ and $w_2 \in \hat{w}(c_2)$. Because the profit function $\pi(w; c, \theta)$ in (9) is strictly decreasing in c , we have $\hat{\pi}(c_1) \geq \pi(w_2; c_1, \theta) > \pi(w_2; c_2, \theta) = \hat{\pi}(c_2)$. Hence, the maximum profit function $\hat{\pi}(c)$ is strictly decreasing in c .

Next, by the definitions of w_1 and w_2 , we have

$$\int_{w_1}^{\bar{w}} \theta(w)f(w) dw [\tau(w_1) - c_1] \geq \int_{w_2}^{\bar{w}} \theta(w)f(w) dw [\tau(w_2) - c_1]$$

$$\int_{w_2}^{\bar{w}} \theta(w)f(w) dw [\tau(w_2) - c_2] \geq \int_{w_1}^{\bar{w}} \theta(w)f(w) dw [\tau(w_1) - c_2].$$

Adding these two inequalities yields

$$\int_{w_1}^{w_2} \theta(w)f(w) dw [c_2 - c_1] \geq 0,$$

which says that w_2 must be at least w_1 since $\theta(w) \geq 0$. ■

The profit-maximizing price may not be strictly increasing in cost c , although the maximum profit is strictly decreasing. For some rationing rules, the profit-maximizing prices at two different cost levels may be identical. For example, suppose that the rationing rule specifies that $\theta(w) = 0$ for $w < \tilde{w}$, and $\theta(w) = 1$ for $w > \tilde{w}$. This scheme supplies consumers (at zero price) if and only if their wealth is below a threshold \tilde{w} . At a low cost, the firm already may sell to all available consumers, setting the price at the willingness to pay $\tau(\tilde{w})$. The optimal price will not reduce further even when cost falls. Figure 1 illustrates this. There, two quantity functions are graphed. First, the quantity function $\hat{w}^m(c)$ is the profit maximizing quantity function when the firm may sell to all consumers, while the quantity function $\hat{w}(c)$ maximizes profit when consumers with wealth less than \tilde{w} are supplied by the public sector. The quantity function $\hat{w}(c)$ coincides

with $\hat{w}^m(c)$ for cost levels above a threshold, and it becomes the horizontal, dotted line when the cost falls below that threshold.

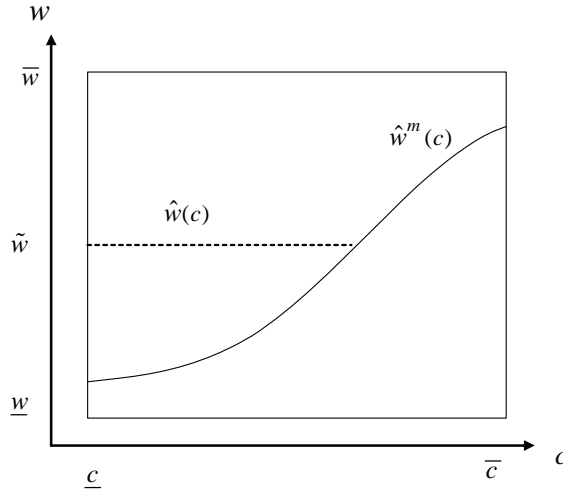


Figure 1: Quantity functions $\hat{w}^m(c)$ and $\hat{w}(c)$.

For some rationing rules, there may be multiple points that maximize profit. For example, suppose that $\theta(w) = 0$ for $w \in [w_1, w_2]$ where $\underline{w} < w_1 < w_2 < \bar{w}$, and $\theta(w) = 1$ otherwise. The public sector supplies only to consumers with wealth levels in a medium range. Figure 2 illustrates the density of consumers available to the private firm. The profit maximizing quantity function is illustrated in Figure 3. For $c < c_1$ or $c > c_2$, the profit-maximizing quantity is unique. For $c \in (c_1, c_2)$, the price remains constant. As the cost falls below c_2 , the firm does not lower its price because all consumers with wealth in $[w_1, w_2]$ are supplied by the public sector. At cost c_1 , the firm makes equal amounts of profit whether it charges a price equal to $\tau(w_2)$ or a price $\tau(w_0)$. The profit from selling to consumers with w between w_0 and w_1 and those with $w > w_2$ at a lower price $\tau(w_0)$ is exactly the same as selling only to those with wealth above w_2 at the higher price $\tau(w_2)$. Notice that the value of w_0 must be strictly below w_1 .

A selection from the correspondence $\hat{w}(c)$ need not be continuous. Nevertheless, because it must be increasing, any point of discontinuity of $\hat{w}(c)$ must be an upward jump, as in Figure 3. An equilibrium is a selection of the profit-maximizing quantity correspondence. By a slight abuse of notation, we denote such a selection by $\hat{w} : [\underline{c}, \bar{c}] \rightarrow [\underline{w}, \bar{w}]$. From Lemma 2, we only need to consider those quantity functions

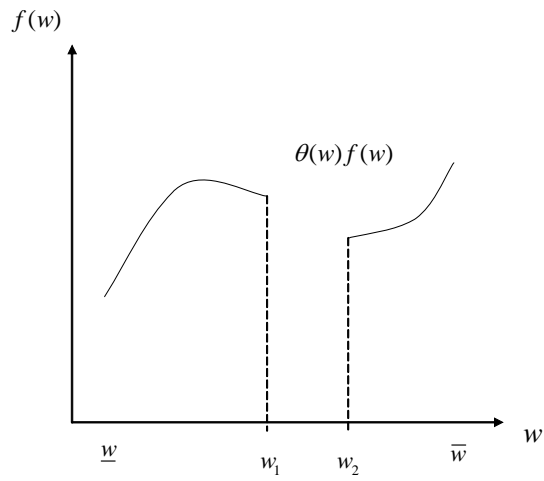


Figure 2: Consumer density under rationing scheme θ .

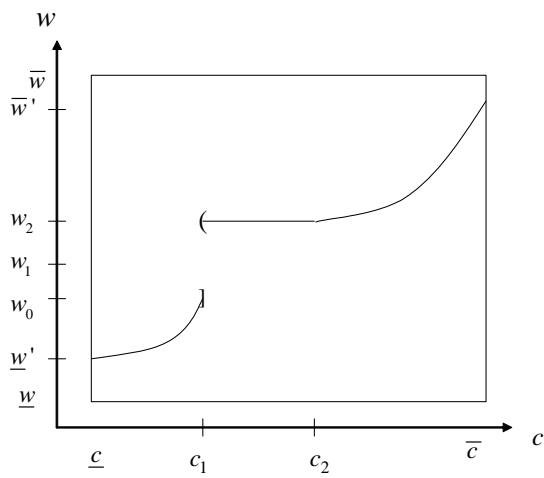


Figure 3: Quantity function $\hat{w}(c)$.

$\hat{w} : [\underline{c}, \bar{c}] \rightarrow [\underline{w}, \bar{w}]$ that are increasing.

For a given quantity function $\hat{w}(c)$ and the corresponding price function $\tau(\hat{w}(c))$, consumer (w, c) buys from the private firm if and only if $w \geq \hat{w}(c)$. Let this set of consumers be denoted by $\Omega \equiv \{(w, c) : w \geq \hat{w}(c)\}$. In Figure 3, this is the set above the graph of $\hat{w}(c)$. If we integrate the utilities of consumers in Ω , we obtain the total consumer benefit.

It is more convenient to view the set Ω as one that is indexed by a function $\hat{c} : [\underline{w}, \bar{w}] \rightarrow [\underline{c}, \bar{c}]$ that is like an “inverse” of \hat{w} . Define $\hat{c}(w) = \sup \{c : w \geq \hat{w}(c)\}$; if there is no $c \in [\underline{c}, \bar{c}]$ such that $w \geq \hat{w}(c)$, set $\hat{c}(w) = \underline{c}$. While the function \hat{w} gives the wealth of the marginal consumer in terms of his cost, the function \hat{c} gives the threshold cost level below which a consumer with wealth w will buy from the firm at price $\tau(\hat{w}(c))$. Whenever \hat{w} is strictly increasing and continuous, the function \hat{c} is its inverse. When \hat{w} is constant, then \hat{c} exhibits a discontinuity. When \hat{w} is discontinuous, then \hat{c} becomes constant. (See Figure 3.) Clearly $\hat{c}(w)$ is increasing in w . The set $\Omega' \equiv \{(w, c) : c \leq \hat{c}(w)\}$ differs from Ω at most for a set of measure zero. Functions \hat{w} and \hat{c} are two equivalent ways to keep track of consumer types who purchase from the private firm.

Given a quantity function \hat{w} (and its equivalent \hat{c}), if the public supplier chooses a rationing scheme θ , the welfare index $V(\theta)$ is

$$\int_{\underline{w}}^{\bar{w}} \theta(w) f(w) \left[\int_{\underline{c}}^{\hat{c}(w)} \{U(w - \tau(\hat{w}(c))) + 1\} g(c) \, dc \right. \\ \left. + \int_{\hat{c}(w)}^{\bar{c}} U(w) g(c) \, dc \right] dw + \int_{\underline{w}}^{\bar{w}} [1 - \theta(w)] f(w) [U(w) + 1] \, dw \quad (12)$$

In this expression, the first term inside the square brackets is the utility of consumers who buy from the private firm while the other terms refer to utilities of consumers who are either given the good for free or refuse to buy from the private firm after having been rationed out. Consumer (w, c) pays the price $\tau(\hat{w}(c))$ when he buys from the private firm, and this price is always lower than his willingness to pay $\tau(w)$. The welfare index can be simplified to

$$V(\theta) = \int_{\underline{w}}^{\bar{w}} \theta(w) f(w) \left[\int_{\underline{c}}^{\hat{c}(w)} \{U(w - \tau(\hat{w}(c))) + 1 - U(w)\} g(c) \, dc \right] dw + \int_{\underline{w}}^{\bar{w}} [U(w) + (1 - \theta(w))] f(w) \, dw, \quad (13)$$

where the first term is the expected inframarginal surplus consumers enjoy when they purchase from the private market, and the second term is the consumer surplus from the public supply.

4 Equilibrium rationing and prices when rationing is based on wealth

An equilibrium is a pair of rationing and quantity schemes (θ, \hat{w}) such that θ maximizes the welfare index (12) subject to the budget constraint (6) given a quantity scheme \hat{w} , and \hat{w} maximizes profit (9) for every c given θ ; that is, the rationing and pricing schemes are mutual best responses against each other.

We let the supplier pick the net density of consumers that will be made available to the private firm θf , and impose the requirement that $0 \leq \theta f \leq f$. The welfare index (13) is linear in θf , and for each w its first-order derivative with respect to θf is

$$\begin{aligned} \frac{\partial V}{\partial \theta f} &= \int_{\underline{c}}^{\hat{c}(w)} \{U(w - \tau(\hat{w}(c))) + 1\} g(c) \, dc + \int_{\hat{c}(w)}^{\bar{c}} U(w) g(c) \, dc - [U(w) + 1] \\ &= \int_{\underline{c}}^{\hat{c}(w)} \{U(w - \tau(\hat{w}(c))) + 1 - U(w)\} g(c) \, dc - 1. \end{aligned} \quad (14)$$

This expression measures the welfare tradeoff between one unit of public provision and letting one unit of consumer to the private market. In the private market, consumer (w, c) faces the price $\tau(\hat{w}(c))$. He will buy from the private market at the price $\tau(\hat{w}(c))$ if his wealth w is above $\hat{w}(c)$. The term inside the integral in (14) is the expected incremental surplus from such transactions for a consumer with wealth w . Against this, the welfare index is reduced by 1, the incremental utility of the good if it is supplied by the public sector at zero charge. The following key lemma establishes a monotonicity in the supplier's preferences.

Lemma 3 *The first-order derivative $\frac{\partial V}{\partial \theta f}$ (14) is increasing in w . It is strictly increasing in $w \in [w_1, w_2]$ unless $\hat{c}(w) = \underline{c}$ for each such w .*

Proof of Lemma 3: Consider w_1 and w_2 with $w_1 < w_2$. Evaluating (14) at w_1 and w_2 and then taking the difference, we have

$$\begin{aligned} \frac{\partial V}{\partial \theta f} \Big|_{w=w_2} - \frac{\partial V}{\partial \theta f} \Big|_{w=w_1} &= \int_{\underline{c}}^{\hat{c}(w_1)} \{[U(w_2 - \tau(\hat{w}(c))) - U(w_1 - \tau(\hat{w}(c)))] - [U(w_2) - U(w_1)]\} g(c) \, dc \\ &\quad + \int_{\hat{c}(w_1)}^{\hat{c}(w_2)} \{U(w_2 - \tau(\hat{w}(c))) + 1 - U(w_2)\} g(c) \, dc \geq 0 \end{aligned} \quad (15)$$

The inequality in (15) follows from the concavity of U , and \hat{c} being increasing. Finally, (15) is zero if and only if $\hat{c}(w_1) = \hat{c}(w_2) = \underline{c}$. ■

Lemma 3 says that the supplier's tradeoff between supplying the good to a consumer and rationing him favors rationing as the consumer's wealth level increases. This is a fundamental principle in our model. The price in the private sector depends only on cost c . Consumer (w, c) gets more surplus from a trade at price $\tau(\hat{w}(c))$ as w increases: $U(w - \tau(\hat{w}(c))) + 1 - U(w)$ is increasing in w .

The last part of Lemma 3 says that the derivative (14) at w is constant if and only if consumers with wealth lower than w do not purchase from the private market. Again, if consumer (w, c) gets to purchase from the private market for some levels of cost, then as w increases, the incremental surplus increases. The derivative (14) must be strictly positive. Figure 4 illustrates a situation where the quantity function \hat{w} becomes constant as cost falls below \tilde{c} , and the price remains at $\tau(\tilde{w})$. Consumers with wealth below \tilde{w} do not buy, and the integral in (15) vanishes.

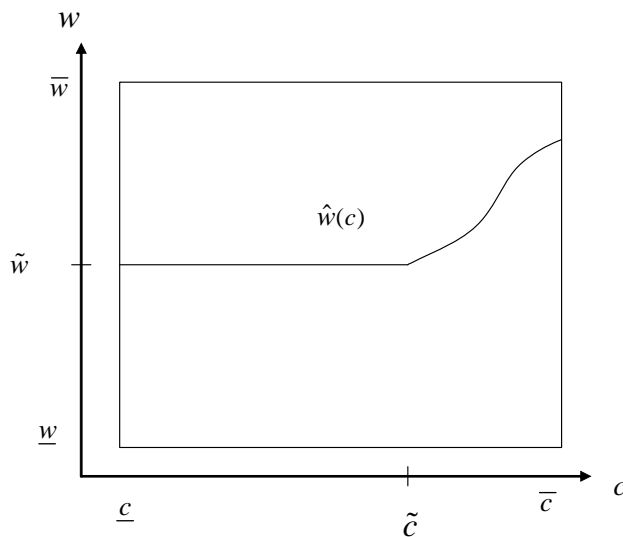


Figure 4: Equilibrium quantity function.

Against a quantity function $\hat{w}(c)$ (and the corresponding $\hat{c}(w)$), the public supplier chooses θf to maximize (12) subject to the budget constraint (6). Using pointwise optimization, we consider the Lagrangean

$$\theta(w)f(w) \left[\int_{\underline{c}}^{\hat{c}(w)} \{U(w - \tau(\hat{w}(c))) + 1\} g(c) \, dc \right] + [1 - \theta(w)]f(w)[U(w) + 1] - \lambda[\gamma(1 - \theta(x))f(x) - B],$$

$$+ \int_{\hat{c}(w)}^{\bar{c}} U(w)g(c) \, dc$$

where λ is the multiplier. The first-order derivative of the Lagrangean with respect to θf is

$$\frac{\partial V}{\partial \theta f} + \lambda \gamma = \int_{\underline{c}}^{\widehat{c}(w)} \{U(w - \tau(\widehat{w}(c))) + 1 - U(w)\} g(c) \, dc - 1 + \lambda \gamma. \quad (16)$$

From Lemma 3, the first-order derivative of the Lagrangean is increasing in w , and increasing strictly whenever some consumers with wealth less than w purchase from the private market.

Lemma 4 *In an equilibrium, the public sector rations consumers with wealth above a threshold \widetilde{w} . That is, in an equilibrium there is $\widetilde{w} < \bar{w}$ such that $\theta(w) = 1$ for $w > \widetilde{w}$.*

Proof of Lemma 4: Because of the limited budget, the public supplier must leave some consumers to the private sector, so that $\theta(w) > 0$ for some w . Because the rationing rule depends only on wealth, some consumers must purchase from the private market. Let $\widetilde{w} = \inf\{w : \theta(w) > 0\}$, and for some $w > \widetilde{w}$, $\widehat{c}(w)$ must be higher than \underline{c} . By Lemma 3, the first-order derivative of the Lagrangean with respect to θf must be strictly increasing in w for $w > \widetilde{w}$. If $\theta(w) > 0$, then the first-order derivative (16) must be nonnegative, and for any $w' > w$, the value of (16) must be strictly positive, and $\theta(w') = 1$. ■

Lemma 4 follows from the monotonicity of the supplier's preferences. If it is optimal for the supplier to ration a consumer at some wealth level, then it must also be optimal to ration all consumers with higher wealth. In any equilibrium, there must exist \widetilde{w} such that $\theta(w) = 1$ if $w > \widetilde{w}$. Lemma 4 does not assert that there is a unique equilibrium. Nor does it say that in an equilibrium, $\theta(w) < 1$ for all $w < \widetilde{w}$.

Lemma 5 *In an equilibrium, the private firm sets a constant price for costs below a threshold \widetilde{c} . That is, $\widehat{w}(c)$ is constant for $c < \widetilde{c}$.*

Proof of Lemma 5: If $\widehat{w}(c)$ is an equilibrium quantity function, it is increasing. Suppose that the Lemma is false. That is, suppose that for some $\widetilde{c} > \underline{c}$, $\widehat{w}(c)$ is strictly increasing for all $c < \widetilde{c}$. Then $\underline{w} < \widehat{w}(c) < \bar{w}$, for $c < \widetilde{c}$. which implies that $\widehat{c}(w) > \underline{c}$ for all w . By Lemma 3, the first-order derivative (16) must be strictly increasing at any w . This means that the set of w at which the first-order derivative (16) vanishes is a single point. Hence, in this equilibrium, the public supplies the good to consumers with wealth below a certain threshold, and $\theta(w) = 0$ for $w < \widetilde{w}$. Against such a rationing scheme, the quantity function $\widehat{w}(c)$ that is

strictly increasing for $c < \tilde{c}$ cannot be optimal, because switching $\hat{w}(c) < \tilde{w}$ to $\hat{w}(c) = \tilde{w}$ yields strictly higher profit. ■

We already know that an equilibrium quantity function is increasing. Lemma 5 says that it cannot be strictly increasing as cost decreases. The reason is that such a strictly increasing quantity function will make the public supplier allocate the good to consumers with low wealth. Nevertheless, a strictly increasing quantity function when costs are low is not a best response against that. When there are no consumers with low wealth and hence low willingness to pay to accept the price reduction, the private firm will leave the price unchanged even as cost falls. Figure 4 illustrates an equilibrium quantity function. Symmetric to Lemma 4, Lemma 5 does not assert that there is a unique equilibrium cost threshold.

The last two lemmas establish the form of an equilibrium. The public sector must ration consumers with high wealth, and the private firm must not sell to consumers with wealth below a threshold. The basic economic principle is the following. Because of the limited budget and the higher surplus for wealthy consumers in the private market, rationing of wealthy consumers must be in an equilibrium. The public sector supplying the less wealthy consumers makes these consumers unavailable to the private firm. As cost decreases, the private firm also has a “corner” solution in its profit-maximizing choice of prices. Prices will become constant even as costs drop further since there may be so few (even zero) consumers with lower willingness to pay to take any price reduction.

We now present an equilibrium.

Proposition 1 *The following is an equilibrium. The public supplier rations all consumers with wealth above a threshold \tilde{w}^s and supplies all consumers with wealth below \tilde{w}^s : $\theta(w) = 1, w > \tilde{w}^s$ and $\theta(w) = 0, w < \tilde{w}^s$. The value of \tilde{w}^s exhausts the budget and is given by $F(\tilde{w}^s)\gamma = B$. The private firm sets a price equal to $\tau(\hat{w}^m(c))$ for $c > \tilde{c}^s$ where $\hat{w}^m(\tilde{c}^s) = \tilde{w}^s$, and a price equal to $\tau(\tilde{w}^s)$ for $c < \tilde{c}^s$.*

Proof of Proposition 1: We verify that the strategies in the proposition form an equilibrium. Given the public supplier’s rationing scheme, it is optimal for the firm to set prices at $\tau(\hat{w}^m(c))$ when $c > \tilde{c}^s$, and at $\tau(\tilde{w}^s)$ when $c < \tilde{c}^s$. Given this quantity function \hat{w} , we have the equivalent function \hat{c} where $\hat{c}(w) = \underline{c}$ for $w < \tilde{w}^s$, and $\hat{c}(w) > \underline{c}$ for $w > \tilde{w}^s$. For (16) we set the multiplier λ to $1/\gamma$. Then the first-order derivative

(16) is zero for $w < \tilde{w}^s$ and strictly positive for $w > \tilde{w}^s$. Moreover, the budget constraint holds as an equality. Hence the rationing scheme defined in the proposition is optimal. The strategies form an equilibrium. ■

The equilibrium in Proposition 1 is illustrated in Figure 4 (set \tilde{c} to \tilde{c}^s and \tilde{w} to \tilde{w}^s there). The private firm sets its prices like it is the monopoly in the market except that it has no access to consumers with wealth below \tilde{w}^s , so the prices in the monopoly quantity schedule $\hat{w}^m(c)$ will stop falling at $\tau(\tilde{w}^s)$ even as cost falls below \tilde{c} . The supplier allocates the good to all consumers with wealth below \tilde{w} and leaves all consumers with higher wealth to purchase from the private market.

This equilibrium is similar to many practical schemes in which poor consumers receive subsidies and free supplies from the government while the rich do not. The reason behind this equilibrium in our model, however, is not one of equity concern. The supplier select among consumers with different wealth levels to participate in the private market. Prices in the private market, however, are dependent on cost. Because of higher willingness to pay, wealthy consumers will realize bigger gains in trade in the private market, and this is the basis for the equilibrium scheme. The private market fully anticipates this, so even when costs decrease, prices may stop falling because consumers with low willingness to pay have already been supplied by the public sector.

Surprisingly, the equilibrium in Proposition 1 is not the only equilibrium in this information regime. Lemmas 4 and 5 do not pin down the supplier's or the firm's strategies. A pricing schedule that becomes constant for low costs may still be consistent with the public supplier rationing some consumers with very low wealth levels. If only a small mass of poor consumers are available to the private market, the private firm will ignore them. To sell to these consumers requires a big price reduction.

The following is another equilibrium in the simultaneous-move game. Let $\epsilon > 0$ and $\delta > 0$ be both small numbers. The supplier's strategies is the following rationing scheme:

$$\begin{aligned}\theta(w) &= 1 \text{ for } \underline{w} < w < \underline{w} + \epsilon \\ \theta(w) &= 0 \text{ for } \underline{w} + \epsilon < w < \tilde{w}^s + \delta \\ \theta(w) &= 1 \text{ for } \tilde{w}^s + \delta < w < \bar{w},\end{aligned}$$

where \tilde{w}^s is defined in Proposition 1. In this rationing rule the supplier shifts some resources from those

with wealth just above the lowest value \underline{w} to those consumers with wealth just above \tilde{w}^s , the equilibrium threshold in Proposition 1. The values of ϵ and δ can be so chosen that the new rationing scheme satisfies the budget. Against this rationing scheme, the private firm sets a quantity function equal to $\hat{w}^m(c)$ for $c > \tilde{c}^s + \eta$ and $\hat{w}^m(\tilde{c}^s + \eta)$ for $c < \tilde{c}^s + \eta$, where \tilde{c}^s is the cost threshold in Proposition 1, and $\eta > 0$ satisfies $\hat{w}^m(\tilde{c}^s + \eta) = \tilde{w}^s + \delta$.

In this equilibrium, the supplier gives the good to some consumers with wealth slightly higher than \tilde{w}^s , but it rations consumers with wealth close to the lowest level. Rationing a small mass of consumers with wealth near \underline{w} would not encourage the private firm to reduce price even when cost becomes very low. Furthermore, because consumers with wealth slightly higher than \tilde{w}^s are now supplied by the public, the private firm's price will not fall all the way to $\tau(\tilde{w}^s)$. Both Lemmas 4 and 5 continue to hold. Figure 4 can still be used to depict this new equilibrium by setting \tilde{c} to $\tilde{c}^s + \eta$ and \tilde{w} to $\tilde{w}^s + \delta$.

Infinitely many equilibria can be constructed in a similar fashion. As long as the private firm does not find it profit-maximizing to reduce price in order to sell to consumers with low willingness to pay, a quantity function like the one in Figure 4 remains a best response. In all these equilibria the public supplier rations some consumers with low wealth, but must ration all consumers with wealth above a threshold.

The equilibrium in Proposition 1 is focal. This is the one that achieves the highest welfare index for the supplier. This is because it has the widest range of price reduction as cost decreases. Any equilibrium different from the one in Proposition 1 would have fewer transactions in the private market, as the next Proposition shows.

Proposition 2 *The equilibrium in Proposition 1 achieves the highest equilibrium welfare for the supplier. In any other equilibrium, the public supplier sets $\theta(w) = 1$, for $w > \tilde{w}^e$, where $\tilde{w}^e > \tilde{w}^s$ (defined by $F(\tilde{w}^s)\gamma = B$ as in Proposition 1) and the firm sets a price equal to $\tau(\hat{w}^m(c))$ for $c > \tilde{c}^e$, and a price equal to $\tau(\tilde{w}^e)$ for $c < \tilde{c}^e$ where $\hat{w}^m(\tilde{c}^e) = \tilde{w}^e$ (with $\tilde{c}^e > \tilde{c}^s$).*

Proof of Proposition 2: In any equilibrium, the budget constraint $\gamma \int_{\underline{w}}^{\bar{w}} (1 - \theta(x))f(x)dx \leq B$ must hold as an equality. The equilibrium in Proposition 1 supplies those with wealth between \underline{w} and \tilde{w}^s , where $F(\tilde{w}^s)\gamma = B$, so $\theta(w) = 0$ for $w < \tilde{w}^s$ and $\theta(w) = 1$ for all $w > \tilde{w}^s$. Consider any other equilibrium. Here,

the supplier must ration some consumers with wealth below \tilde{w}^s because the budget constraint must hold. Hence, for this equilibrium the threshold \tilde{w}^e at which $\theta(w) = 1$ for all $w > \tilde{w}^e$ must be strictly higher than \tilde{w}^s .

Let \tilde{c}^e be defined by $\hat{w}^m(\tilde{c}^e) = \tilde{w}^e$. Clearly, $\tilde{c}^e > \tilde{c}^s$ because \hat{w}^m is strictly increasing and $\tilde{w}^e > \tilde{w}^s$. We now show that the firm's equilibrium price is $\tau(\hat{w}^m(\tilde{c}^e))$ for $c < \tilde{c}^e$. At $w < \tilde{w}^e$, the first-order derivative of the Lagrangean (16) must be nonnegative; if that derivative was negative, then $\theta(w) = 0$, which would violate the budget constraint. Because $\theta(w) = 1$ for all $w > \tilde{w}^e$, the value of (16) is positive for all $w > \tilde{w}^e$. Therefore, by Lemma 3, for $w < \tilde{w}^e$ the value of (16) must be exactly zero and $\hat{c}(w) = \underline{c}$ for $w < \tilde{w}^e$. Because $\theta(w) = 1$ for all $w > \tilde{w}^e$ the equilibrium quantity must be $\hat{w}^m(c)$ for $c > \tilde{c}^e$, and remains constant at $\hat{w}^m(\tilde{c}^e)$ for $c < \tilde{c}^e$.

From $\tilde{c}^s > \tilde{c}^e$ and $\tilde{w}^e > \tilde{w}^s$, by comparing the values of (12) across the equilibria in Proposition 1 and the alternative, we conclude that the supplier's payoff is higher in the equilibrium in Proposition 1. ■

We have assumed a monopolistic private sector. The extension to an imperfectly competitive sector poses no conceptual problem. For our model of a homogeneous good, we consider a Cournot model. Let there be N firms in the private sector. Given a rationing scheme θ , let each firm choose a quantity function $\hat{q}_i(c)$, where $i = 1, \dots, N$. The total supply is $q(c) = \sum_{i=1}^N q_i(c)$. For the market to clear the marginal consumer is $\hat{w}(c)$ where $\int_{\hat{w}(c)}^{\bar{w}} \theta(w)f(w) dw = q(c)$, and the price in the private sector is $\tau(\hat{w}(c))$. All results derived above continue to hold for any given number of firms in the private sector.

Next, we can extend our model to the case of a perfectly competitive private sector. Here, we simply let the price in the private sector be marginal cost: $\tau(c) = c$. Given this pricing function, the corresponding quantity function $\hat{w}(c)$ is implicitly defined by $U(\hat{w} - c) + 1 = U(\hat{w})$. Lemma 3 can be applied on this quantity function. Because the perfectly competitive quantity function is strictly increasing, the derivative (16) is strictly increasing for all values of w . Lemma 4 continues to hold. In sum we have the following result.

Corollary 1 *If the private market is perfectly competitive so that prices there are equal to marginal costs,*

the public sector uses the entire budget on consumers with low wealth levels: $\theta(w) = 0$, for $w < \tilde{w}^s$, and $\theta(w) = 1$, for $w > \tilde{w}^s$ where $F(\tilde{w}^s)\gamma = B$.

5 Equilibrium rationing and prices when rationing is based on wealth and cost

In this section we let the supplier observe both wealth and cost information; the private firm continues to observe only consumers' costs. A rationing policy is $\phi : [\underline{w}, \bar{w}] \times [\underline{c}, \bar{c}] \rightarrow [0, 1]$. Suppose that the firm observes that a consumer's cost of service is c , the density of consumer available to the private firm is $\phi(w, c)f(w)$. If it sets a price equal to $\tau(w)$, the total mass of consumers purchasing is $\int_w^{\bar{w}} \phi(w, c)f(w)dw$, and the profit is

$$\int_w^{\bar{w}} \phi(w, c)f(w)dw [\tau(w) - c]. \quad (17)$$

We use the same notation and let $\hat{w}(c)$ be the quantity that maximizes profit (17), and $\hat{\pi}(c)$ be the maximum profit.⁵

Consider an equilibrium, a rationing function $\phi : [\underline{w}, \bar{w}] \times [\underline{c}, \bar{c}] \rightarrow [0, 1]$ and a quantity function $\hat{w} : [\underline{c}, \bar{c}] \rightarrow [\underline{w}, \bar{w}]$. Given the quantity function, consumer (w, c) buys from the private firm if and only if $U(w - \tau(\hat{w}(c))) + 1 \geq U(w)$. Therefore, when the supplier rations consumer (w, c) , that consumer obtains a utility $\max[U(w - \tau(\hat{w}(c))) + 1, U(w)]$ from the private sector. The welfare index from a policy ϕ is

$$\int_{\underline{c}}^{\bar{c}} \int_{\underline{w}}^{\bar{w}} \{\phi(w, c) \max[U(w - \tau(\hat{w}(c))) + 1, U(w)] + [1 - \phi(w, c)] [U(w) + 1]\} f(w)g(c) dw dc \quad (18)$$

Our first result reports that in equilibrium, the supplier does not supply to consumers who have high costs, irrespective of their wealth levels.

Lemma 6 *In an equilibrium, consumer (w, c) is rationed when his cost is above a threshold, $\tilde{c} > \underline{c}$. That is, $\phi(w, c) = 1$ for $c > \tilde{c} > \underline{c}$ and any w .*

Proof of Lemma 6: In an equilibrium, the public supplier chooses ϕ to maximize (18) subject to the budget constraint (8), given a quantity function $\hat{w}(c)$. Consider pointwise maximization at each (w, c) . The

⁵Given a rationing rule, the profit function is continuous in w , so the quantity $\hat{w}(c)$ that maximizes it is well-defined.

first-order derivative with respect to ϕ is

$$\max [U(w - \tau(\hat{w}(c))) + 1, U(w)] - [U(w) + 1] + \lambda c, \quad (19)$$

where $\lambda > 0$ is the multiplier.

Given a quantity function, if consumer (w, c) prefers to purchase from the private firm, the expression in (19) becomes

$$U(w - \tau(\hat{w}(c))) - U(w) + \lambda c \quad (20)$$

Otherwise, the expression in (19) becomes

$$-1 + \lambda c. \quad (21)$$

When $U(w - \tau(\hat{w}(c))) + 1 > U(w)$, the value in (20) is larger than (21).

Consider all consumers (w, c) who prefer not to purchase from the private firm. Now the first-order derivative is given by (21), which is strictly increasing in c . Furthermore, if $-1 + \lambda c > 0$, then the expression in (20) is positive.

Now we claim that in an equilibrium expression (21) must not be always strictly positive. Suppose not, then $-1 + \lambda c > 0$ for all c , and therefore, the first-order derivative (19) is always strictly positive. The supplier rations all consumers so that $\phi(w, c) = 1$ for all w and c . This implies that the supplier does not use its budget, and this cannot be optimal. We conclude that there must exist $\tilde{c} > \underline{c}$ such that (21) vanishes at $c = \tilde{c}$.

Finally, whenever $c > \tilde{c}$, $-1 + \lambda c > 0$ so that the first-order derivative (19) is strictly positive. We conclude that $\phi(w, c) = 1$, all $c > \tilde{c}$ and any w . ■

When the supplier observes wealth and cost information, a standard cost-effectiveness analysis applies. The basic principle says that if the cost is too high relative to the benefit, the consumer should not be given the good. Now the availability of the private supply for consumers with high cost does not alter this principle. When the cost is sufficiently high, consumers who do not purchase from the private market do not get allocated the good. For those consumers who choose to purchase, they must obtain a higher surplus than those who do not. Therefore, the supplier continues to ration them. Lemma 6 says that for high cost

consumers the rationing rule is the same as the optimal allocation without a private sector.

Now we characterize the equilibrium rationing scheme for low cost consumers. Consider now $c < \tilde{c}$.

Lemma 7 *In an equilibrium, consumer (w, c) is given the good when his cost is below a threshold. That is, $\phi(w, c) = 0$ for $\underline{c} < c < \tilde{c}$, any w .*

Proof of Lemma 7: Consider $c < \tilde{c}$. From Lemma 6, the value of (21) is strictly negative for $c < \tilde{c}$. For $w < \hat{w}(c)$, $U(w - \tau(\hat{w}(c))) + 1 < U(w)$ so that (19) takes the value of (21) and it is negative. We conclude that $\phi(w, c) = 0$.

If $\hat{w}(c) = \bar{w}$, the proof is complete, so now we prove that in an equilibrium, for $c < \tilde{c}$, $\hat{w}(c) = \bar{w}$. Suppose not; that is, suppose that $\hat{w}(c) < \bar{w}$. Now at $w = \hat{w}(c)$, $U(w - \tau(\hat{w}(c))) + 1 = U(w)$, so that the derivative (19) is negative. For $\epsilon > 0$ and sufficiently small, $U(w + \epsilon - \tau(\hat{w}(c))) - U(w + \epsilon) + \lambda c < 0$, so that the derivative (19) remains negative, and $\phi(w + \epsilon, c) = 0$. Now given that at c , consumers with wealth between $w = \hat{w}(c)$ and $w = \hat{w}(c) + \epsilon$ are supplied by the public sector, the private firm will raise the price from $\tau(\hat{w}(c))$, or equivalently raise the equilibrium quantity function from $\hat{w}(c)$. This contradicts the assumption that $\hat{w}(c) < \bar{w}$ is an equilibrium quantity function. We conclude that for $c < \tilde{c}$, $\hat{w}(c) = \bar{w}$. ■

Due to cost effectiveness, the supplier will assign the good to low-cost consumers who do not purchase from the private sector. Now some wealthy consumers with low cost may obtain a higher surplus from the private sector if the price is low. Lemma 7 says that this cannot happen in an equilibrium. Consider the marginal consumer (w, c) ; he pays a price $\tau(w)$ from the private market and obtains a zero incremental surplus. Now because his cost is low, the public supplier prefers to allocate the good to him, giving him a positive incremental surplus. By continuity, the supplier also assigns the good to those with wealth slightly above w ; this assignment, however, eliminates these consumers from the private market. The best response by the private firm is to raise the price and increase the wealth of the marginal consumer. This unravelling argument continues until the private firm raises the price to $\tau(\bar{w})$, and does not sell to consumers with low cost. As a result, in an equilibrium, the price is too high even for low-cost consumers, and the supplier provides the good to all these consumers.

The proofs of Lemma 6 and Lemma 7 make minimal assumptions on the structure of prices set by the private firm. A simpler set of arguments suffices for equilibrium characterization in the regime where rationing is based on wealth and cost than rationing based on wealth alone. To summarize, we present the following.

Proposition 3 *If the public sector rations consumers based on wealth and cost information, the equilibrium rationing function is identical to the optimal rationing function when the private sector is inactive. That is, $\phi(w, c) = 0$ for $c < c^s$ and $\phi(w, c) = 1$, otherwise. The private firm chooses the monopoly quantity scheme $\hat{w}^m(c)$ for $c > c^s$, and $\hat{w}^m(c) = \bar{w}$ for $c < c^s$.*

Equilibria in the two information regimes are very different. When rationing is based on wealth, in the focal equilibrium in Proposition 1 consumers with low wealth levels get the good, but when cost information is also available, the equilibrium rationing scheme becomes independent of wealth, and assigns the good to low-cost consumers. In both information regimes, the market is segmented, and selection occurs in the private market with the private firm selling only to those rationed consumers who have higher willingness to pay.

There is not a clear welfare comparison between equilibria across the two information regimes. Although the equilibria in Propositions 1 and 3 are simple to describe, they refer to rationing of very different consumer types. In both regimes, equilibrium prices in the private market always follow the monopoly schedule \hat{w}^m for costs that are high. When public rationing is based on wealth, the price there becomes constant when costs are low. The extent of the fall in prices depends on consumers' willingness to pay function $\tau(w)$ as well as the distribution of wealth $F(w)$.

Clearly, the results in Proposition 3 apply directly to imperfectly competitive firms operating in the private market. Equilibrium rationing when the private market is perfectly competitive may be a little different. Here, some low-cost and wealthy consumers may be rationed. When the private market is competitive, consumers face marginal cost pricing, so prices are given by $\tau(c) = c$. The supplier's welfare index is

$$\int_{\underline{c}}^{\bar{c}} \int_{\underline{w}}^{\bar{w}} \{\phi(w, c) \max [U(w - c) + 1, U(w)] + [1 - \phi(w, c)] [U(w) + 1]\} f(w)g(c) dw dc. \quad (22)$$

Corollary 2 *Suppose the private market is competitive so that prices there are equal to marginal costs. The public sector rations all consumers with cost above a threshold. For those consumers with costs below the threshold, the public sector supplies a consumer if and only if his wealth is below a value determined by his cost. That is, $\phi(w, c) = 1$ for $c > c'$, some $c' > \underline{c}$ and any w ; $\phi(w, c) = 1$ for $c < c'$ and $w > \mu(c)$ where the function μ is implicitly defined by $U(\mu - c) - U(\mu) + \lambda c = 0$ for a constant $\lambda > 0$; otherwise, $\phi(w, c) = 0$.*

Proof of Corollary 2: First, set $\tau(\widehat{w}(c))$ to c in the proof of Lemma 6. It follows that $\phi(w, c) = 1$ for $c > c'$, some $c' > \underline{c}$. Now consider $c < c'$. For those consumer (w, c) who do not purchase from the private sector, the proof of Lemma 7 applies and $\phi(w, c) = 0$.

Consider consumer (w, c) , $c < c'$, and $U(w - c) + 1 > U(w)$, so that this consumer purchases from the private sector at cost c . The first-order derivative of the Lagrangean with respect to ϕ is

$$U(w - c) - U(w) + \lambda c, \tag{23}$$

where $\lambda > 0$ is the multiplier for the budget constraint. Expression (23) is negative at $c < c'$ and $w = \widehat{w}(c)$. Hence if there exists $w' > \widehat{w}(c)$ such that $U(w' - c) - U(w') + \lambda c > 0$, then $\phi(w', c) = 1$. The function $w'(c)$ in the proposition is implicitly defined by $U(w - c) - U(w) + \lambda c = 0$, for $c < c'$. ■

Corollary 2 is illustrated in Figure ?? The upward sloping line $\widehat{w}(c)$ is the marginal consumer given marginal cost pricing. The line above $\widehat{w}(c)$ is the function μ defined in the Corollary.⁶ Consumers with cost $c < c'$ and wealth between $\widehat{w}(c)$ and $\mu(c)$ strictly prefer to purchase the good at cost than to go without, yet the supplier will assign the good to them for free. Very wealthy consumers are rationed even when their costs are low.

For consumers with costs higher than c' , it is not cost effective to supply them. The logic in Lemma 6 applies. For those with costs lower than c' , the logic in Lemma 7 applies with one modification. Cost consideration alone warrants the allocation. Nevertheless, wealthy consumers obtain higher incremental surplus from the private market. The condition in Corollary 2 can be rewritten as

$$U(\mu - c) + 1 = U(\mu) + 1 - \lambda c.$$

⁶By definition, $U(\widehat{w}(c) - c) + 1 = U(\widehat{w}(c))$. Now because $c < c'$, we have $1 > \lambda c$. It follows that the value of μ that satisfies $U(\mu - c) + 1 = U(\mu) + (1 - \lambda c)$ must be greater than $\widehat{w}(c)$.

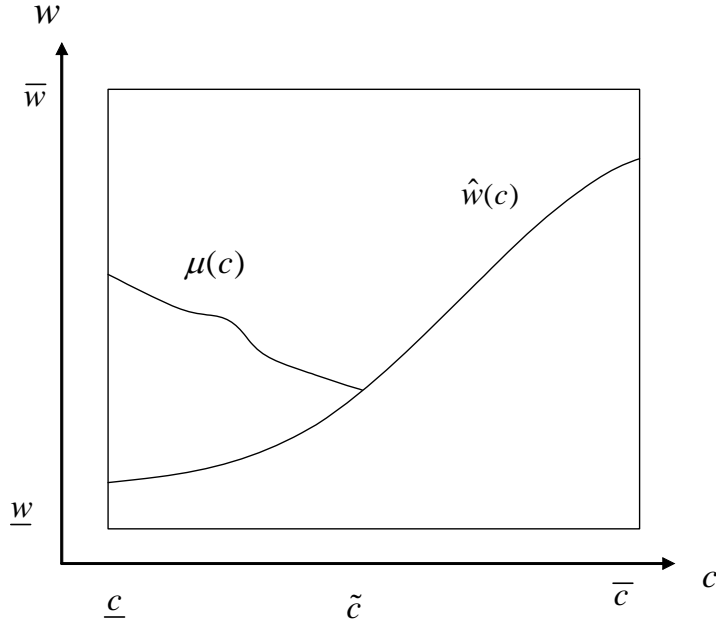


Figure 5: Optimal rationing under a competitive private market

The left-hand side expression is the utility of a consumer with wealth μ buying from the private market; the right-hand side expression is the social net benefit, where λ is the multiplier of the budget constraint and λc measures the utility equivalent of cost c . If $w > \mu(c)$, the consumer gets more utility from the private market than the social net benefit from the good, so will be rationed. In contrast to Lemma 7, there is no unravelling of equilibrium price best responses in the private market.

6 Alternative assumptions on preferences

In this section, we discuss two alternative assumptions on consumer preferences. We have used a separable utility function. If we have used the general utility specification, the willingness-to-pay function is implicitly defined by $U(w - \tau, 1) = U(w, 0)$. From total differentiation, we have

$$\frac{d\tau}{dw} = 1 - \frac{\frac{\partial U}{\partial w}(w, 0)}{\frac{\partial U}{\partial w}(w - \tau, 1)}.$$

If we further have assumed that wealth and the good are complements in that the partial derivative of U with respect to w is larger with consumption of the good than without ($\frac{\partial U}{\partial w}(w, 1) \geq \frac{\partial U}{\partial w}(w, 0)$), the willingness-to-pay function is increasing. If the goods are substitutes ($\frac{\partial U}{\partial w}(w, 1) \leq \frac{\partial U}{\partial w}(w, 0)$), the willingness-to pay

function may still be increasing if the degree of substitutability is small.

Consider first the rationing based on wealth. For the general utility function, the welfare index is $\int_{\underline{w}}^{\bar{w}} [1 - \theta(w)]U(w, 1) dF + \int_{\underline{w}}^{\bar{w}} \theta(w)U(w, 0) dF$. We can derive the optimal rationing rule when the private market is inactive, so we consider the maximization of the welfare index subject to the budget constraint. The Lagrangean is $[1 - \theta]U(w, 1) + \theta U(w, 0) + \lambda[B - (1 - \theta)\gamma]$. From pointwise maximization, the first-order derivative with respect to θ is $-U(w, 1) + U(w, 0) + \lambda$. This first-order derivative is increasing in w if and only if wealth and the good are substitutes. In this case, the public sector supplies the good to consumers with low wealth, so that $\theta(w) = 0$ if and only if w is below a threshold. Results from our model with a separable utility function for the consumer will be robust if this is the case.

Next, consider rationing based on wealth and cost. We can again write down the welfare index, and derive the optimal rationing scheme when the private sector is inactive. By pointwise optimization of the welfare index subject to the budget constraint, the Lagrangean is $[1 - \phi]U(w, 1) + \phi U(w, 0) + \lambda[B - (1 - \phi)c]$, and its first-order derivative with respect to ϕ is $-U(w, 1) + U(w, 0) + \lambda c$, which is strictly increasing in c , and strictly increasing in w if and only if $\frac{\partial U}{\partial w}(w, 1) \leq \frac{\partial U}{\partial w}(w, 0)$. Define (w^*, c^*) by $-U(w^*, 1) + U(w^*, 0) + \lambda c^* = 0$. If wealth and the good are substitutes, the optimal rationing rule sets $\phi(w, c) = 0$ if and only if $w \leq w^*$ and $c \leq c^*$. If they are complements, the optimal rationing rule sets $\phi(w, c) = 0$ if and only if $w \geq w^*$ and $c \leq c^*$. In either case, the optimal cost-effectiveness rule assigns the good to consumers whose cost of provision is low. It is to be modified due to the change in the marginal utility of wealth from the consumption of the good. Given these tendencies, we believe that the results under the separable utility function will be quite robust.

Now we return to the separable utility function, but allow the consumer's benefit to vary with c . Let $\beta(c)$ be the benefit from the consumption of the good for consumer (w, c) . When the public supplier does not observe cost, only the expected value of $\beta(c)$ can be used in computing optimal rationing policies. Our results regarding the public supplier's response against a private market quantity function remains valid.

Now suppose that the supplier has both wealth and cost information. The standard cost-effectiveness comparison will be modified as follows. If the private market is inactive, the comparison is between $\beta(c)$

and the shadow cost of c : $\beta(c) - \lambda c$, where λ is the multiplier of the budget constraint. When $\beta(c)$ is an increasing and concave function, then the optimal rationing rule will still assign the good to those consumers where $\beta(c) - \lambda c > 0$. In other words, it remains true that those consumers with cost below a threshold will be assigned.

The private firm's profit-maximizing strategy is more complicated. The willingness to pay τ for consumer (w, c) is given by $U(w - \tau) + \beta(c) = U(w)$. Obviously τ depends on w and c , and is increasing in both arguments. At cost c , if the firm sets a price p , the demand is $\int_{p < \tau(w, c)} dF(w)$ and the profit is $[p - c] \int_{p < \tau(w, c)} dF(w)$. Profits may well be increasing in c , and prices may not be monotone in cost. When high-cost consumers benefit a lot from the good, then the firm gets a higher profit selling to them. Selection takes a different form in this case. There may be a natural segmentation of the market: the public sector's optimal rationing rule will assign the good to those with costs below a threshold, while the private market caters for those with high costs.

7 Concluding remarks

To be continued

References

- [1] Arrow, K., "An Utilitarian Approach to the Concept of Equality in Public Expenditure," *Quarterly Journal of Economics*, Vol. 85, No. 3, (1971), pp. 409-415.
- [2] Barros, P. and P. Olivella, "Waiting Lists and Patient Selection," *Journal of Economics & Management Strategy*, Vol. 14(3), (2005), pp.623-646, 09.
- [3] Besley, T. and Coate S., "Public Provision of Private Goods and the Redistribution of Income", *The American Economic Review*, Vol. 81, No. 4. (1991), pp. 979-984.
- [4] Bergstrom, T., "Soldier of Fortune," in *Essays in Honor of K.J. Arrow*, ed. Walter Heller and Ross Starr, (1986).
- [5] Blackorby, C. and Donaldson, D., "Cash versus Kind, Self-Selection, and Efficient Transfers", *American Economic Review*, Vol. 78, No. 4. (1988), pp. 691-700.
- [6] Boadway, R., and Marchand, M., "The use of Public Expenditures for Redistributive Purposes," *Oxford Economic Papers*, New Series, Vol. 47, No. 1., (1995), pp. 45-59.
- [7] De Fraja, G., "The Design of Optimal Education Policies," *The Review of Economic Studies*, Vol. 69, (2002), pp. 437-466.
- [8] Ellis, R., "Creaming, skimping and dumping: provider competition on the intensive and extensive margins", *Journal of Health Economics*, Vol. 17, Issue 5, (1998), pp. 537-555
- [9] Hoel, Michael, "What should (public) health insurance cover?" *Journal of Health Economics*, Vol. 26 (2007) pp. 251-262.
- [10] Mirrlees, J. A., "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, Vol. 38, (1971), No. 114, pp. 175-208.