



# EUI Working Papers

MWP 2009/30  
**MAX WEBER PROGRAMME**

OPTIMAL PUBLIC RATIONING: PRICE RESPONSE  
AND COST EFFECTIVENESS

Simona Grassi Ching-to Albert Ma



EUROPEAN UNIVERSITY INSTITUTE, FLORENCE  
MAX WEBER PROGRAMME

*Optimal Public Rationing:  
Price Response and Cost Effectiveness*

SIMONA GRASSI CHING-TO ALBERT MA

This text may be downloaded for personal research purposes only. Any additional reproduction for other purposes, whether in hard copy or electronically, requires the consent of the author(s), editor(s). If cited or quoted, reference should be made to the full name of the author(s), editor(s), the title, the working paper or other series, the year, and the publisher.

The author(s)/editor(s) should inform the Max Weber Programme of the EUI if the paper is to be published elsewhere, and should also assume responsibility for any consequent obligation(s).

ISSN 1830-7728

© 2009 Simona Grassi Ching-to Albert Ma

Printed in Italy  
European University Institute  
Badia Fiesolana  
I – 50014 San Domenico di Fiesole (FI)  
Italy  
[www.eui.eu](http://www.eui.eu)  
[cadmus.eui.eu](http://cadmus.eui.eu)

## **Abstract**

We study optimal public rationing of an indivisible good and private sector price response. Consumers differ in their wealth and cost of provisions. Due to a limited budget, some consumers must be rationed. Public rationing determines the characteristics of consumers who seek supply from the private sector, where a firm sets prices based on consumer cost information and in response to the rationing rule. We consider two information regimes. In the first, the public supplier rations consumers according to their wealth information. In equilibrium, the public supplier must ration both rich and poor consumers. Supplying all poor consumers would leave only rich consumers in the private market, and the firm would react by setting a high price. Rationing some poor consumers is optimal, and implements price reduction in the private market. In the second information regime, the public supplier rations consumers according to consumer wealth and cost information. In equilibrium, rationing is based on cost-effectiveness and consumers are allocated the good if and only if their costs are below a threshold.

## **Keywords**

Publicly provided private goods, Publicly provided goods: mixed markets, Government expenditures and welfare programs, Analysis of health care markets, Government policy; provision and effects of welfare programs

*Simona Grassi*  
*Max Weber Fellow, 2008-2009*  
*European University Institute*  
*Villa La Fonte, Via Delle Fontanelle, 10*  
*I - 50014 San Domenico di Fiesole (FI)*  
*Simona.Grassi@eui.eu*

*Ching-to Albert Ma*  
*Department of Economics*  
*Boston University*  
*270 Bay State Road*  
*Boston, MA 02215*  
*ma@bu.edu*



# 1 Introduction

Rationing of services and goods such as health, education and housing by a public supplier is very common. Free public supply often coexists with active private markets. Rationing and public supply are two sides of the same action. When the public supplier provides a good for free to some consumers, rationed consumers may consider purchasing from the private sector at their own expense. Rationing therefore determines the portfolio of consumers in the private market. A private firm must recognize this, and choose its pricing strategy accordingly. In this paper we study how rationing affects prices in these private markets and characterize the optimal rationing policy.

The analysis of how the market reacts to a public program is crucial to the design of an optimal policy. It is well documented in the public finance literature that public programs often crowd out purchases in the private market. Consider the health sector. Since the seminal paper by Gruber and Cutler (1996), many studies have tried to measure the extent to which the expansion of public health programs targeted towards the less wealthy patients, such as Medicaid, crowds out private insurance coverage. In the last twenty years in the US, the number of uninsured but also publicly insured individuals has risen (see Gruber and Simon, 2008). Despite the lack of consensus about the magnitude of the crowding-out effect, the majority of studies agree that part of the rise in the number of publicly insured individuals can be explained by the fall in the number of privately insured ones. A policy aimed at reducing the number of the uninsured cannot ignore the impact on and the reaction of the private market.

A reactive private market cannot be modeled as a perfectly competitive market. Suppose that a public supplier provides for free to some eligible consumers a given quantity of an indivisible good, which is also traded in a perfectly competitive market. If the supply curve is not perfectly inelastic, the leftward shift in the demand curve would lead to a lower equilibrium price and quantity. However, the equilibrium quantity falls by less than the increase in the quantity supplied for free by the public supplier. Therefore, the total number of consumers who get the good after the policy is implemented necessarily increases. Crowding out cannot be explained by the assumption that the firms in the market are price takers. Our model provides a general framework for studying the design of an optimal rationing policy when the private market is modeled

as a monopoly. It could be used to understand the mechanisms behind the crowding-out of private purchases due to the expansion of free public provision in the health sector. Cutler and Gruber [1996] suggests some "mechanisms through which employer-provided coverage could fall as Medicaid eligibility increases". Among these, they consider a price (premium) rise in the private market. More precisely, they suggest that employers may reduce their contribution to insurance premiums, or that "as Medicaid coverage rises, providers may maintain their incomes by charging higher prices to privately insured patients, lowering the demand for private insurance ." A price-reactive private market is hence consistent with a plausible explanation of the crowding out effect.

In our model, consumers are heterogenous in two dimensions. The costs of providing the good to them differ and they have different wealth levels. Consumers wealth heterogeneity is natural, and more wealthy consumers may be more willing to pay for services. In the health market, cost heterogeneity is common. Patients with higher severity levels are more costly to treat. This is also true in education. The cost of educating a student depends on the student's aptitude and other demographic factors. We consider the supply of an indivisible good, and further standardize the unit so that it represents one unit of improvement in well-being. We then assume that its benefit is the same to all consumers. As an example, consider a hip replacement procedure. Treating patients with higher levels of disability requires high costs. A hip replacement will allow a consumer to walk about without pain, which we regard as a unit of health improvement. Consumers who are more wealthy may be more willing to pay for the hip replacement. Achieving literacy can be thought of as the good to be provided to students, and we normalize it so that literacy represents a unit of benefit.

We consider the effect of rationing on prices under two information regimes. In the first, the public supplier observes consumer' wealth level and can credibly commit to a rationing scheme based on consumer wealth. In the second, the public supplier observes both consumer' wealth and cost levels, and credibly commits to a rationing scheme based on both pieces of information.

The public supplier has access to wealth information through tax returns. It may well have access to and use cost information. This is our second information regime. We also consider the case where the

public supplier is unable to use cost information in our first information regime. 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.

As already mentioned above, we model the private market as a monopoly. The firm does not observe consumers' wealth levels, but it does observe their cost characteristic and knows the rationing scheme implemented by the public supplier. The firm chooses a price function based on cost.

Our model is unlike the typical regulation model, where the regulated, private firm has more information than a social planner. In our model, a public supplier may even have more information than the private firm. The critical issue we consider is how the allocation of the public supplier's limited budget for consumers affects the private firm's price responses.

Rationing defines which consumers are entitled to public provision. Suppose that rationing is based on wealth, and poor consumers are supplied by the public while rich consumers are rationed. Only the rich consumers will consider the private market. Anticipating that the consumers available in the market are rich and have a high willingness to pay, a firm tends to raise prices. High prices will still be optimal even if the costs of providing to these rich consumers are low. This is recognized as a form of cream-skimming.

In the health economics literature, cream-skimming usually refers to providers selecting of low-cost, hence, profitable patients under a prospective payment system. This conception ignores the possibility that consumers have different willingness to pay, and this is important in a mixed system where the private market is active. A high-cost, rich consumer may well be more profitable than a low-cost, poor consumer. The price-cost margin for a rich consumer may be larger. In our model, cream-skimming means that the firm can price discriminate according to the cost of provision and also that the firm uses information on the supplier's rationing policy to deduce whether the consumer on the market has high or low willingness to pay.

The public supplier can mitigate cream-skimming by using wealth information and by choosing the proportion of poor and rich consumers in the market. If the public supplier rations some poor consumers, making them available to the private market, a private firm may then find it attractive to lower prices when consumers have good risks or low costs. The mixture of rich and poor consumers in the private market makes

price reduction in the private sector an equilibrium response. This is the main result of the analysis under the first information regime. Using total consumer utility as a welfare index, we show that in equilibrium, the public supplier must ration both poor and rich consumers, and implement some price reduction in the private sector.

Those concerned with equity may find rationing the poor disagreeable. Nevertheless, we show that rationing some poor consumers will yield a first-order gain due to price reduction. Supplying *all* poor consumers will exacerbate cream-skimming, eliminating this gain. Unless social preferences are so extreme that only the welfare of poor consumers matters, the price reduction due to the rationing of *some* poor consumers will be beneficial. However, we extend the analysis to consider the optimal policy of a public supplier who has equity concerns and include welfare weights in the social welfare function.

In the second information regime, rationing is based on both wealth and cost information. Surprisingly, in equilibrium the public supplier ignores the wealth information and rations consumers according to *cost-effectiveness*. All consumers with costs below a threshold are supplied, while those with higher costs are rationed. Since the rationing scheme is independent of wealth, the firm cannot anticipate the wealth composition of consumers in the market. The firm cannot do better than setting prices as if the public sector did not exist, although it must only sell to consumers with higher costs. The public supplier induces more price reduction when rationing is based on wealth than when rationing is based on wealth and cost. We analyze the effect of the introduction of welfare weights in this second information regime too.

In Grassi and Ma (2008) we study a similar model, but the public rationing and private price schemes are chosen simultaneously. That model offers a longer term perspective on the interaction, because rationing and price schemes must be mutual best responses. We find that when the public supplier observes consumers' wealth and cost levels, it implements a rationing policy based on cost-effectiveness, the same as in the model here. Under rationing based on wealth and cost information, equilibria are robust with respect to the public supplier's commitment power.

If rationing is only based on wealth, the game in Grassi and Ma (2008) has multiple equilibria, all of which differ from the equilibrium here. While the public supplier rationing the poor to induce price reduction is the

equilibrium strategy in the sequential game, price reduction is never implemented in the simultaneous-move game.

The literature on rationing and the private sector usually assumes an exogenous supply in the private market. Barros and Olivella (2005) consider doctors who self-refer patients in the public sector to the private sector. Prices paid by patients in the private sector are fixed, while doctors only refer low-cost patients. Iversen (1997) studies waiting-time rationing when there is a private market. Hoel and Sæther (2003) consider the effect of competitive supplementary insurance on a national health insurance system. Hoel (2007) derives the optimal cost-effectiveness rule when patients have access to a competitive private market. A competitive private market has been a common assumption.. This may be relevant in many settings, but is unlikely to be true always. In fact, when the private market pricing rule is fixed, one only can study how the private market influences public policies. By contrast, we study the case where public policies influence private market responses.

Strategic interaction between the public and the private sectors is studied in the literature of mixed oligopolies. A mixed oligopoly is defined as a market structure where a public enterprise coexists with one or more profit-maximizing firms. See Cremer et al. (1991), DeFraja and Delbono (1990), Merrill and Schneider (1966), Beato and Mas Colell (1984). Works in the mixed oligopoly literature use a variety of assumptions on whether goods are homogeneous or differentiated, whether the public firm has a first-mover advantage, and whether the public firm has some budget available in order to offer goods at below costs. This literature focuses on the markets for public utilities, such as telecommunication, transportation, water, and energy, where mixed oligopolies are common. A mixed oligopoly is also common in the health care, education and housing sectors. Issues of rationing and selection are especially relevant in the health care sector, but have not been the focus of that literature. Our research connects between the analysis of mixed oligopoly and rationing and selection.

The paper is organized as follows. Section 2 lays out the model. Section 3 and its subsections describe the firm's choice of the profit maximizing prices in the continuation equilibrium and the equilibrium rationing when the public supplier observes only consumers' wealth level. Section 4 and its subsections focus on the

information regime where wealth and cost levels are observed by the public supplier. The last Section draws some conclusions. An Appendix contains all proofs.

## 2 The model

There is a set of consumers. Each consumer's wealth is either  $w_1$  or  $w_2$ , with  $w_1 < w_2$ . Let  $m_i > 0$  be the mass of consumers with wealth  $w_i$ ,  $i = 1, 2$ . We call consumers with wealth  $w_1$  poor consumers, and consumers with wealth  $w_2$  rich consumers.

Each consumer may consume at most one unit of an indivisible good. If a consumer pays a price  $p$  for the good, his utility is  $U(w_i - p) + 1$ , while if he does not consume the good (and pays 0), his utility is  $U(w_i)$ . The function  $U$  is strictly increasing and strictly concave. The good gives a unit utility increment to a consumer. We can use a general utility function where the utilities from consuming the good at price  $p$  and from not consuming the good are  $U(w - p, 1)$  and  $U(w, 0)$ , respectively. A separable utility function simplifies the analysis.

A consumer's willingness to pay for the good,  $\tau_i$ , is defined by the following:

$$U(w_i - \tau_i) + 1 = U(w_i), \quad i = 1, 2, \tag{1}$$

so  $\tau_i$  is the maximum price a consumer with wealth  $w_i$  is willing to pay. Because  $U$  is strictly concave,  $\tau_1 < \tau_2$ ; a rich consumer is more willing to pay for the good than a poor consumer. Similarly, by the strict concavity of  $U$ , for any  $p > 0$ , we have  $U(w_2) - U(w_2 - p) < U(w_1) - U(w_1 - p)$ . It follows that a rich consumer enjoys more surplus than a poor consumer when the good is sold at price  $p$ :

$$U(w_1 - p) + 1 - U(w_1) < U(w_2 - p) + 1 - U(w_2).$$

The cost of providing the good to a consumer is random. Let  $c$  denote this cost,  $G : [\underline{c}, \bar{c}] \rightarrow [0, 1]$  its distribution function, and  $g$  the corresponding density, both defined on a positive support, and with  $g > 0$ . The distribution  $G$  is independent of wealth and uncorrelated to the utility benefit from consuming the good, which is fixed at 1. In the health care example, this means that patients with different severity levels obtain the same incremental utility from the good. Furthermore, we assume that the hazard rate  $G(c)/g(c)$

is increasing. Let  $\gamma$  be the expected value of  $c$ . We assume that  $\underline{c} < \tau_1 < \bar{c} \leq \tau_2$ . The last, weak inequality involves no loss of generality because a firm attempting to sell to any consumer with cost higher than  $\tau_2$  will not make any sale.

A public supplier has a budget  $B$  which is insufficient to provide the good for free to all consumers, so we assume  $B < (m_1 + m_2)\gamma$ . We consider two information regimes. In the first, the public supplier can use a non-price rationing mechanism based on wealth. In the second, the public supplier uses a non-price rationing mechanism based on both wealth and cost. There is also a private market, and we model it as a monopoly. The private firm does not observe consumers' wealth but does observe their costs.

When rationing is based on consumers' wealth, a rationing policy is given by  $(\theta_1, \theta_2)$ ,  $0 \leq \theta_i \leq 1$ ,  $i = 1, 2$ . For each wealth class  $w_i$ , the regulator rations  $\theta_i m_i$  consumers, and supplies  $(1 - \theta_i)m_i$  consumers. When rationing is based on consumers' wealth and costs, a rationing policy is given by  $(\phi_1, \phi_2)$ , a pair of functions  $\phi_i(c) : [\underline{c}, \bar{c}] \rightarrow [0, 1]$ ,  $i = 1, 2$ . The value  $\phi_i(c)g(c)$  is the density of consumers with wealth  $w_i$  and cost  $c$  who are rationed. For consumers with  $w_i$ , the mass of rationed consumers with cost less than  $c$  is

$$m_i \int_{\underline{c}}^c \phi_i(x)g(x)dx,$$

and the mass of supplied consumers with cost less than  $c$  is

$$m_i \int_{\underline{c}}^c [1 - \phi_i(x)]g(x)dx.$$

Consumers who are rationed may consider buying from the private market. The public supplier's payoff is the sum of consumer utilities.

There is a private market and we model it as a monopoly. The firm observes a consumer's cost  $c$ , but not his wealth  $w_i$ . In our model, if the firm managed to observe both consumers' wealth and cost, it would extract all consumer surplus. Given the public supplier's rationing policy, the private firm chooses prices as a function of costs to maximize profits.

We study the subgame-perfect equilibria of the following extensive-form games:

**Stage 0:** Nature determines that a mass  $m_i$  of consumers have wealth  $w_i$ ,  $i = 1, 2$ , and draws a cost realization for each consumer according to the distribution  $G$ . The private firm observes a consumer's

cost realization, but not his wealth. Under rationing based on wealth, the public supplier observes a consumer's wealth, but not the cost realization. Under rationing based on wealth and cost, the public supplier observes a consumer's wealth and cost.

**Stage 1:** Under rationing based on wealth, the public supplier sets a rationing policy  $(\theta_1, \theta_2)$ ,  $0 \leq \theta_i \leq 1$ , supplying  $(1 - \theta_i)m_i$  consumers with wealth  $w_i$ ,  $i = 1, 2$ . Under rationing based on wealth and cost, the public supplier sets a rationing policy  $(\phi_1, \phi_2)$ ,  $\phi_i(c) : [\underline{c}, \bar{c}] \rightarrow [0, 1]$ , supplying  $[1 - \phi_i(c)]m_i$  consumers with wealth  $w_i$  and cost  $c$ .

**Stage 2:** The private firm sets a price for each cost realization.

**Stage 3:** Consumers who are rationed by the public supplier may purchase from the private firm at the prices set at Stage 2.

### 3 Equilibrium rationing and prices in wealth-based rationing

#### 3.1 Continuation equilibrium prices

For a given rationing policy  $(\theta_1, \theta_2)$ , we derive the firm's continuation equilibrium prices in Stage 2. Because consumers are either poor or rich, the firm will set its price to either  $\tau_1$  or  $\tau_2$ . Clearly at any cost above  $\tau_1$ , the firm must set the price at  $\tau_2$ , selling only to rich consumers. Suppose that the cost  $c$  decreases below  $\tau_1$ . The firm may set a low price  $\tau_1$ , selling to both rich and poor consumers, or a high price  $\tau_2$ , selling only to rich consumers; these profits are respectively:

$$\pi(\tau_1; c \leq \tau_1) \equiv (m_1\theta_1 + m_2\theta_2)[\tau_1 - c] \quad (2)$$

$$\pi(\tau_2; c \leq \tau_1) \equiv m_2\theta_2[\tau_2 - c]. \quad (3)$$

The profits in (2) and (3) are linear in  $c$ , and (2) decreases in  $c$  at a faster rate than (3); these functions are illustrated in Figure 1. For some rationing policies  $(\theta_1, \theta_2)$  there may be a cost level  $c_1 < \tau_1$  in  $[\underline{c}, \bar{c}]$  such that  $\pi(\tau_1; c_1) = \pi(\tau_2; c_1)$ , as in Figure (1). This value of  $c_1$  is given by

$$c_1 \equiv \tau_1 - \frac{m_2\theta_2}{m_1\theta_1}(\tau_2 - \tau_1). \quad (4)$$

As the cost drops below  $\tau_1$ , a price reduction is worthwhile only if there are enough poor consumers relative to rich ones. The value of  $c_1$  is the cost threshold at which a price reduction occurs and is decreasing in the ratio of available rich to poor consumers,  $\frac{m_2\theta_2}{m_1\theta_1}$ . If there are few poor consumers in the market, the cost has to be much lower than  $\tau_1$  for a price reduction to occur. In an extreme, if only the rich consumers are rationed and all the poor are supplied by the public sector, the firm will not reduce the price at all. We summarize the firm's continuation equilibrium prices by the following:

**Lemma 1** *Given a rationing scheme  $(\theta_1, \theta_2)$ , in a continuation equilibrium if  $c_1$  in (4) is greater than  $\underline{c}$ , the firm sets the high price  $\tau_2$  if  $c > c_1$ , and the low price  $\tau_1$  if  $c < c_1$ ; if  $c_1$  in (4) is less than  $\underline{c}$ , the firm always sets the high price  $\tau_2$ .*

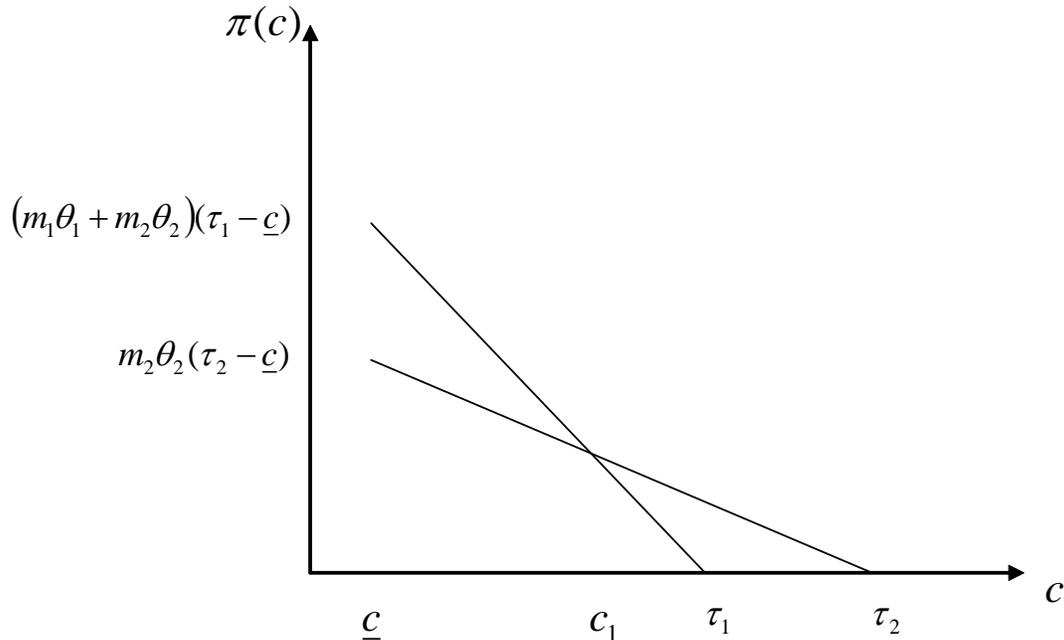


Figure 1: Comparison of profit between setting high and low prices

We should rule out any situation where the cost has no influence on prices when the budget is zero. For this, we assume that when all consumers are rationed because  $B = 0$ , the firm switches from low to high price at an interior cost threshold. That is, at  $\theta_1 = \theta_2 = 1$  in (4) there is  $c_m$  in the interior of  $[\underline{c}, \bar{c}]$ , where

$$c_m = \tau_1 - \frac{m_2}{m_1}(\tau_2 - \tau_1). \quad (5)$$

### 3.2 Equilibrium rationing

We introduce a new notation  $\beta \equiv B/\gamma$ . Because  $B$  denotes the budget available to the public supplier and  $\gamma$  the expected cost, the term  $\beta$  is the number of consumers to whom the public supplier can provide the good. Given the continuation equilibrium prices, the total consumer utility is:

$$m_1 \left[ (1 - \theta_1) (U(w_1) + 1) + \theta_1 \left\{ \int_{\underline{c}}^{c_1} [U(w_1 - \tau_1) + 1] dG + \int_{c_1}^{\bar{c}} U(w_1) dG \right\} \right] +$$

$$m_2 \left[ (1 - \theta_2) (U(w_2) + 1) + \theta_2 \left\{ \int_{\underline{c}}^{c_1} [U(w_2 - \tau_1) + 1] dG + \int_{c_1}^{\bar{c}} [U(w_2 - \tau_2) + 1] dG \right\} \right].$$

In this expression, terms involving  $(1 - \theta_i)$  are consumers' utilities when they receive the public supply at no charge. Terms involving  $\theta_i$  are the market outcomes. For poor consumers, if their costs are below  $c_1$ , they purchase at  $\tau_1$ , which actually leaves them no surplus (see definition of  $\tau_i$  in (1)). Similarly, for rich consumers, if their costs are above  $c_1$ , they purchase at price  $\tau_2$ , earning no surplus. However, if rich consumers' costs are below  $c_1$ , they earn a surplus  $U(w_2 - \tau_1) + 1 - U(w_2) > 0$ , since the price  $\tau_1$  is lower than their willingness to pay  $\tau_2$ .

Using  $U(w_i - \tau_i) + 1 = U(w_i)$ ,  $i = 1, 2$ , we simplify the previous expression to

$$[m_1 U(w_1) + m_2 U(w_2) + m_1(1 - \theta_1) + m_2(1 - \theta_2)] + m_2 \theta_2 \int_{\underline{c}}^{c_1} [U(w_2 - \tau_1) + 1 - U(w_2)] g(c) dc, \quad (6)$$

where  $c_1 \geq \underline{c}$  characterizes the firm's continuation equilibrium price strategy. The term  $m_1(1 - \theta_1) + m_2(1 - \theta_2)$  is the total number of consumers receiving public supply, while  $U(w_2 - \tau_1) + 1 - U(w_2)$  is the surplus a rich consumer enjoys when he purchases at price  $\tau_1$ . In an equilibrium the budget  $B$  must be exhausted. Hence, we replace  $m_1(1 - \theta_1) + m_2(1 - \theta_2)$  by  $\beta$ , and simplify (6) to

$$V(\theta_2, c_1) \equiv [m_1 U(w_1) + m_2 U(w_2) + \beta] + m_2 \theta_2 \int_{\underline{c}}^{c_1} [U(w_2 - \tau_1) + 1 - U(w_2)] g(c) dc. \quad (7)$$

An equilibrium is a rationing policy  $(\theta_1, \theta_2)$  and the continuation equilibrium price strategy in Lemma 1  $c_1$  that maximize (7), subject to the cost threshold definition (4), the budget constraint

$$m_1(1 - \theta_1) + m_2(1 - \theta_2) \leq \beta \equiv \frac{B}{\gamma} (< m_1 + m_2), \quad (8)$$

and the boundary conditions  $\underline{c} \leq c_1$ , and  $0 \leq \theta_i \leq 1$ ,  $i = 1, 2$ .

**Proposition 1** *In the equilibrium, the public supplier rations consumers in each wealth class:  $\theta_1 > 0$  and  $\theta_2 > 0$ , while the firm charges the low price  $\tau_1$  when the consumer's cost is below a threshold  $c_1^*$ , where  $\underline{c} < c_1^* < \tau_1$ .*

Proposition 1 (whose proof is in the appendix) says that for any budget, the public supplier must ration some poor consumers and some rich consumers. Only rich consumers potentially gain from the private market, so rationing some rich consumers must be an equilibrium. The gain will not be realized if the firm does not reduce the price. To realize this potential gain, the public supplier must leave enough poor consumers in the market so that the ratio between rich and poor consumers makes it attractive for the private firm to reduce the price. Therefore, the equilibrium cost threshold  $c_1$  at which the price drops from  $\tau_2$  to  $\tau_1$  must be strictly above  $\underline{c}$ . We emphasize that rationing of consumers occurs in equilibrium even if the budget is sufficient to provide for an entire class of consumers. The following characterizes the equilibrium cost threshold  $c_1$ , and the rationing policy  $(\theta_1, \theta_2)$ .

**Proposition 2** *If the budget  $B$  is sufficiently large, the value of  $c_1^*$  is given by the unique solution of*

$$\frac{G(c_1^*)}{g(c_1^*)} = \frac{(\tau_1 - c_1^*)(\tau_2 - c_1^*)}{\tau_2 - \tau_1},$$

*$\theta_1 < 1$ , and  $\theta_2 < 1$ ; the public supplier supplies some consumers in each wealth class. If the budget is small, either  $\theta_1$  or  $\theta_2$  may be equal to 1 and the public supplier may ration an entire wealth class; the budget constraint then can be used to solve for the optimal rationing policy.*

Proposition 2 (whose proof is in the appendix) reports that there are three possible rationing outcomes. In the “interior” solution, the value of  $c_1$  is obtained by the appropriate first-order conditions, and the boundary conditions  $\theta_i \leq 1$  do not bind. In the “corner” solutions, either  $\theta_1 = 1$  or  $\theta_2 = 1$ .

The maximization of the objective function in (7) is equivalent to the maximization of

$$m_2 \theta_2 [U(w_2 - \tau_1) + 1 - U(w_2)] G(c_1).$$

This expression describes the surplus earned by the fraction  $\theta_2$  of rationed rich consumers with cost below the threshold  $c_1$ . The inframarginal gain is constant, and the optimization problem is equivalent to the

even simpler maximization of  $m_2\theta_2G(c_1)$ . This objective function is increasing in both  $\theta_2$  and  $c_1$ , but the equilibrium price and budget constraints, respectively (4) and (8), and the boundary conditions  $\theta_i \leq 1$  limit how high  $\theta_2$  and  $c_1$  can be.

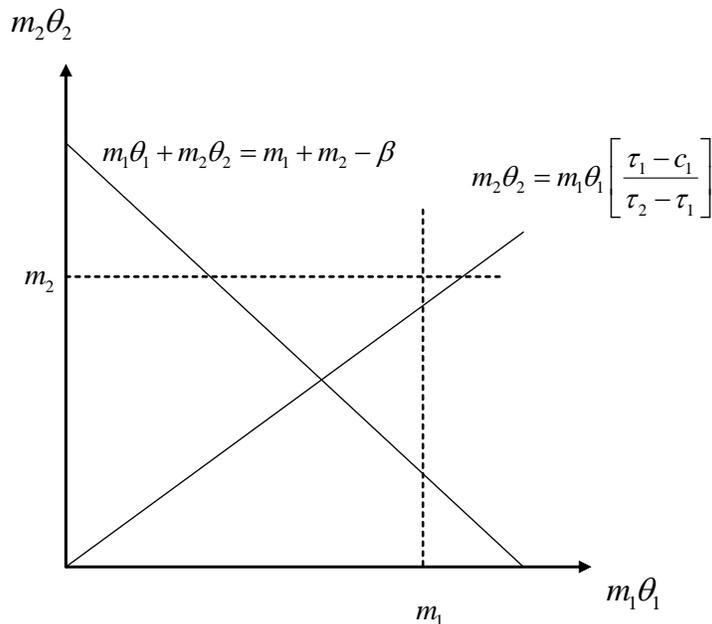


Figure 2: Budget and cost threshold constraints; boundary conditions.

The constraints are graphed in Figure 2. The downward sloping line is the budget constraint (8), while the upward sloping line through the origin is the equilibrium cost threshold (4) for some  $c_1$ ; as the value of  $c_1$  increases, the line becomes flatter. The two dotted lines are the boundary constraints for  $\theta_i$ .

By solving for  $\theta_1$  and  $\theta_2$  with (4) and (8), and then substituting into the objective function, we have the objective function expressed in terms of  $c_1$  alone, and then after ignoring constants, we can write it as  $\frac{\tau_1 - c_1}{\tau_2 - c_1}G(c_1)$ . (Details are in the proof.) The equilibrium value  $c_1^*$  in the Proposition achieves the maximum of this objective function with the boundary conditions  $\theta_i \leq 1$  ignored.

The value of  $c_1^*$  balances the trade-off between rationing more rich consumers and rationing more poor consumers. Rationing a rich consumer allows him to realize a surplus if his cost is below the cost threshold. Rationing a rich consumer, however, implies supplying a poor consumer. With fewer poor consumers in the market, the cost has to fall below a lower  $c_1$  threshold before the private firm reduces its price. This then

reduces the likelihood  $G(c_1)$  that a rich consumer will benefit from the private market.

Clearly,  $c_1^*$  is the equilibrium value when the boundary conditions  $\theta_i \leq 1$  are satisfied. In this case we can use the budget and cost threshold constraints to solve for  $\theta_i$  after setting  $c_1$  to  $c_1^*$ , and both  $\theta_1$  and  $\theta_2$  are smaller than 1:

$$\theta_1 = \frac{m_1 + m_2 - \beta}{m_1 \left[ 1 + \frac{G(c_1^*)}{g(c_1^*)} \frac{1}{\tau_2 - c_1^*} \right]} < 1 \quad \text{and} \quad \theta_2 = \frac{m_1 + m_2 - \beta}{m_2 \left[ 1 + \frac{g(c_1^*)}{G(c_1^*)} (\tau_2 - c_1^*) \right]} < 1. \quad (9)$$

(Details of the computation are in the proof in the Appendix.) Figure 2 illustrates such a case where the intersection of the budget and cost threshold constraints (at  $c_1 = c_1^*$ ) is in the interior of the area bounded by the boundary conditions represented by the two dotted lines.

The boundary conditions are unlikely to bind when the budget is big, so that in Figure 2, the budget line is located closer to the origin. If the budget is small, the budget line is located farther from the origin. One of the two boundary conditions may be violated when  $c_1$  is set at  $c_1^*$ . In this case, the unconstrained maximization of  $\frac{\tau_1 - c_1}{\tau_2 - c_1} G(c_1)$  is infeasible. Instead, the intersection points between the budget line and the boundary conditions must be considered. Either  $\theta_1 = 1$  or  $\theta_2 = 1$ . Then, the value of cost threshold  $c_1$  can be obtained by the constraint (4).

From (9), we can see that the total number of consumers  $m_i$  in a wealth class will likely determine whether that group will be completely rationed when the budget is small. We can also see this in Figure 2. If the value of  $m_1$  is small, then the vertical dotted line is closer to the origin. It is more likely that  $\theta_1$  becomes 1. That is, to implement a cost reduction, the public supplier may have to let the private supplier potentially sell to all poor consumers. Conversely, if  $m_2$  is small, then the horizontal dotted line is closer to the origin. It is more likely that  $\theta_2$  becomes 1. In this case, the public supplier finds it optimal to let all rich consumers gain from trading in the private market.

How do our results change when the public supplier is concerned with equity? If the public supplier is more concerned with the poor, we may let the public supplier's payoff be a weighted sum of poor and rich consumers' utilities, with a weight  $1 + \epsilon$  for poor consumers and 1 for rich consumers, where  $\epsilon \geq 0$ . Formally, we simply replace  $m_1$  in (6) by  $m_1(1 + \epsilon)$ . Then the equilibrium will be given by the solution for

the maximization of this modified objective function subject to the budget and cost-threshold constraints, as well as the boundary conditions.

By the Maximum Theorem, the solution is continuous in  $\epsilon$  so our results, Propositions 1 and 2, are robust: a small increase of  $\epsilon$  from zero will only change the optimal rationing rule slightly, and in the case of corner solutions, may not at all. The first-order condition that yields the solution  $c_1^*$  in Proposition 2 becomes  $\frac{G(c_1)}{g(c_1)} - \frac{\epsilon}{\Delta g(c_1)} = \frac{(\tau_1 - c_1)(\tau_2 - c_1)}{\tau_2 - \tau_1}$  so that the cost threshold at which price reduction occurs tends to be lower. Generally, the concern for equity favors supplying poor consumers, reducing their presence in the private market. Equity concern tends to reduce the likelihood of price reduction.

We conclude this section by describing two benchmarks. The first is when there is no active private market. The second is when the private market is competitive so that prices there are always equal to costs. If there is no active private market, any rationing policy that exhausts the budget is optimal. Each consumer benefits a unit of utility from the good, irrespective of his wealth level. The total increase in expected consumer benefit is simply  $B/\gamma$ , or the total number of consumers who can be served by the budget, and this is independent of how the budget is distributed to the consumers.

If there is a competitive private market, the unique equilibrium will have the budget first used to supply poor consumers. If all poor consumers can be supplied, then the remaining budget will be used to supply a fraction of rich consumers. This formal result can be found in Grassi and Ma (2008) for consumers having a continuum of wealth levels, and it applies here. The intuition is this. Given that prices are always equal to marginal costs, the regulator does not seek to influence pricing decisions in the private market. For any given price in the private market, rich consumers benefit more than poor consumers. By rationing the rich consumers, the public supplier allows more inframarginal gains from trade in the private market.

## 4 Equilibrium rationing and prices in wealth-cost based rationing

### 4.1 Continuation equilibrium prices

Given a rationing policy  $(\phi_1, \phi_2)$ ,  $\phi_i(c) : [\underline{c}, \bar{c}] \rightarrow [0, 1]$ , we derive the continuation equilibria. Again, there are only two possible equilibrium prices in the private market, the low price  $\tau_1$  and the high price  $\tau_2$ .

For any  $c > \tau_1$ , the unique best response by the private firm is  $\tau_2$ . For any  $c$  between  $\underline{c}$  and  $\tau_1$ , the firm chooses between the low price  $\tau_1$  and the high price  $\tau_2$ . The firm's profit from the low price  $\tau_1$  is  $[m_1\phi_1(c) + m_2\phi_2(c)][\tau_1 - c]$ ; the profit is  $m_2\phi_2(c)[\tau_2 - c]$  if the firm sets the high price  $\tau_2$ . Therefore, the firm sets the low price  $\tau_1$  if  $[m_1\phi_1(c) + m_2\phi_2(c)][\tau_1 - c] \geq m_2\phi_2(c)[\tau_2 - c]$ , or

$$m_1\phi_1(c)[\tau_1 - c] \geq m_2\phi_2(c)[\tau_2 - \tau_1]. \quad (10)$$

It sets a high price  $\tau_2$  if (10) is violated, and it may randomize between  $\tau_1$  and  $\tau_2$  if (10) holds as an equality.

These are the continuation equilibrium prices.

We now define a new, indicator function for continuation equilibria when  $c < \tau_1$ . Let  $p : [\underline{c}, \tau_1] \rightarrow [0, 1]$ . Given a policy  $(\phi_1, \phi_2)$ , we set  $p(c) = 1$  if (10) holds,  $p(c) = 0$  if (10) is violated, and  $p(c)$  to a number between 0 and 1 if (10) holds as an equality. The function  $p$  is the probability of price reduction when the cost is between  $c$  and  $\tau_1$ . If  $p(c)$  takes the value 0, we understand it to mean no price reduction, and the private firm chooses the high price  $\tau_2$ , whereas if  $p(c)$  takes the value 1, we understand it to mean a price reduction, and the private firm chooses the low price  $\tau_1$ . If the value of  $p(c)$  is a fraction, the private firm randomizes between the two prices.

**Lemma 2** *For  $c$  between  $\underline{c}$  and  $\tau_1$ , any continuation equilibrium is given by a function  $p : [\underline{c}, \tau_1] \rightarrow [0, 1]$  satisfying the following two inequalities:*

$$p(c) \{m_1\phi_1(c)[\tau_1 - c] - m_2\phi_2(c)[\tau_2 - \tau_1]\} \geq 0 \quad (11)$$

$$[1 - p(c)] \{m_1\phi_1(c)[\tau_1 - c] - m_2\phi_2(c)[\tau_2 - \tau_1]\} \leq 0. \quad (12)$$

Lemma 2 (whose proof is in the appendix) introduces a “complementary” function  $p(c)$  to describe the continuation equilibria. The key to understanding continuation equilibria is whether at cost  $c$  a price reduction will be implemented by the policy  $(\phi_1, \phi_2)$ . Price reduction is a best response if (10) holds. The term inside the curly brackets of (11) and (12) is the difference between the left-hand and right-hand sides of (10). The variable  $p(c)$  is an indicator; it takes the value 1 when there is a price reduction, and 0 otherwise. The inequalities (11) and (12) are complementary conditions that make  $p(c)$  and (10) consistent. By Lemma 2, any equilibrium can be described by  $p : [\underline{c}, \tau_1] \rightarrow [0, 1]$  satisfying (11) and (12).

For ease of exposition, we extend the function  $p$  from the domain  $[\underline{c}, \tau_1]$  to  $[\underline{c}, \bar{c}]$ , and set  $p(c) = 0$  for  $c > \tau_1$ . This simply says that there is no price reduction for  $c > \tau_1$ . This extension allows us to write payoffs in a simpler way.

## 4.2 Equilibrium rationing

Given a rationing policy  $(\phi_1, \phi_2)$ ,  $\phi_i(c) : [\underline{c}, \bar{c}] \rightarrow [0, 1]$ , we have the continuation equilibrium given by Lemma 2. For values of  $c$  higher than  $\tau_1$ , the equilibrium price in the private sector must be  $\tau_2$ , and  $p(c) = 0$ . For values of  $c$  less than  $\tau_1$ , consumers who are rationed may purchase from the private sector at the low price when  $p(c) = 1$  or at the high price when  $p(c) = 0$ . We now write the public supplier's payoff in the continuation equilibrium:

$$\begin{aligned} & \int_{\underline{c}}^{\bar{c}} \{m_1[1 - \phi_1(c)][U(w_1) + 1] + m_2[1 - \phi_2(c)][U(w_2) + 1]\} dG(c) \\ & + \int_{\underline{c}}^{\bar{c}} m_1\phi_1(c) \{[1 - p(c)]U(w_1) + p(c)[U(w_1 - \tau_1) + 1]\} dG(c) \\ & + \int_{\underline{c}}^{\bar{c}} m_2\phi_2(c) \{[1 - p(c)][U(w_2 - \tau_2) + 1] + p(c)[U(w_2 - \tau_1) + 1]\} dG(c). \end{aligned}$$

In this expression, the first integral is the sum of the utilities of all those consumers supplied by the public system; each consumer gets one unit of utility without incurring any cost. In the second integral, we write the sum of the utilities of poor consumers who are rationed. A poor consumer who has cost  $c$  will encounter a price reduction with probability  $p(c)$  (in a continuation equilibrium). If there is not a price reduction, the poor consumer does not buy, so his payoff is  $U(w_1)$ . If there is a price reduction, the poor consumer buys at price  $\tau_1$ , hence the term  $U(w_1 - \tau_1) + 1$ . In the last integral, we write the sum of the utilities of rich consumers who are rationed. If there is not a price reduction, the rich consumer buys at  $\tau_2$ , hence the term  $U(w_2 - \tau_2) + 1$ . If there is a price reduction, he buys at  $\tau_1$ , hence the term  $U(w_2 - \tau_1) + 1$ .

As in the case when rationing is based only on wealth information, the gain in utility when consumers participate in the market is due to the rich consumers purchasing at the low price  $\tau_1$ . Poor consumers either do not buy or buy at their reservation price  $\tau_1$ , gaining no surplus from the private market. This is clearly

seen when we use the definitions of  $\tau_1$  and  $\tau_2$  to simplify the payoff to:

$$\begin{aligned} & m_1 U(w_1) + m_2 U(w_2) + \int_{\underline{c}}^{\bar{c}} \{m_1[1 - \phi_1(c)] + m_2[1 - \phi_2(c)]\} dG(c) \\ & + \int_{\underline{c}}^{\bar{c}} m_2 \phi_2(c) p(c) [U(w_2 - \tau_1) + 1 - U(w_2)] dG(c). \end{aligned} \quad (13)$$

In (13), the terms inside the first integral is the consumers' utility gain from the public supply, and the terms inside the second integral is the incremental gain of rationed rich consumers who purchase in the private market at the low price  $\tau_1$ .

The optimal rationing policy is one that maximizes (13) subject to the budget constraint, and the continuation equilibrium prices in the private market. By Lemma 2, the continuation equilibrium price is given by  $p(c)$  satisfying (11) and (12). Ignoring the constant terms in (13), we write down the maximization program for the public supplier's equilibrium policy. Choose a policy  $(\phi_1, \phi_2)$  and the corresponding price reduction function  $p$  to maximize

$$\int_{\underline{c}}^{\bar{c}} \{m_1[1 - \phi_1(c)] + m_2[1 - \phi_2(c)]\} dG(c) + \int_{\underline{c}}^{\bar{c}} m_2 \phi_2(c) p(c) [U(w_2 - \tau_1) + 1 - U(w_2)] dG(c) \quad (14)$$

subject to

$$B - \int_{\underline{c}}^{\bar{c}} \{m_1[1 - \phi_1(c)] + m_2[1 - \phi_2(c)]\} c dG(c) \geq 0 \quad (15)$$

$$p(c) \{m_1 \phi_1(c) [\tau_1 - c] - m_2 \phi_2(c) [\tau_2 - \tau_1]\} \geq 0 \quad (16)$$

$$[1 - p(c)] \{m_1 \phi_1(c) [\tau_1 - c] - m_2 \phi_2(c) [\tau_2 - \tau_1]\} \leq 0, \quad (17)$$

and the boundary conditions  $0 \leq \phi_i(c)$ ,  $p(c) \leq 1$ ,  $i = 1, 2$ , each  $c$  in  $[\underline{c}, \bar{c}]$ , and  $p(c) = 0$  for  $c > \tau_1$ . Call this Program R. The budget constraint (15) says that the total expected cost from public supply must not exceed the budget. For completeness, we have rewritten the two inequalities in Lemma 2 as (16) and (17).

**Proposition 3** *In the optimal rationing policy based on wealth and cost, the public supplier rations consumers if and only if their costs are above a threshold. That is, in an equilibrium,*

$$\begin{aligned} \phi_1(c) &= \phi_2(c) = 0 \quad \text{for } c < c^B \\ \phi_1(c) &= \phi_2(c) = 1 \quad \text{for } c > c^B, \end{aligned}$$

where the cost threshold  $c^B$  is defined by

$$\int_{\underline{c}}^{c^B} (m_1 + m_2)cdG(c) = B.$$

Figure 3 shows the three cases that make up the proof of Proposition 3 (which is in the appendix). In all cases, the equilibrium rationing rule is based only on cost, not on wealth. Case 1 is where the budget is large, sufficient to supply some consumers with cost above  $\tau_1$ . Here, all consumers in the private market have costs higher than  $\tau_1$ , and the private firm sets the high price  $\tau_2$ . Poor consumers with high costs are not provided with the good by the public supplier; nor are they willing to buy in the private market. Rich consumers with high costs buy from the private market but earn no surplus.

Case 2 is where the budget is of medium size, lower than  $\tau_1$  but higher than  $c_m$ , where  $c_m$ , defined in (5), is the cost level below which the private firm sets the low price  $\tau_1$  when all rich and poor consumers are in the market. Here, as in Case 1, the price in the private market is always  $\tau_2$ . The public supplier continues to provide for low-cost consumers. Even though there are some consumers with cost lower than  $\tau_1$ , the private firm does not sell to them.

Finally, Case 3 is where the budget is small, less than  $c_m$ . Here, the public supplier induces the private firm to charge the low price  $\tau_1$  when the cost is below  $c_m$ . At higher costs, the price is set at  $\tau_2$ . A low price in the private sector is an equilibrium if and only if the public supplier has a small budget.

Before we present the intuition behind Proposition 3, we present a benchmark where the private market is inactive. Here, the public supplier chooses a rationing policy  $(\phi_1, \phi_2)$  to maximize consumer's utility subject to the budget constraint. This is actually a special case of this model: simply set  $p(c) = 0$ , and omit constraints (16) and (17). The policy maximizing (14) (with  $p(c) = 0$ ) subject only to the budget constraint (15) is a simple *cost-effectiveness* rule.<sup>1</sup> Each consumer obtains a unit of utility, independent of his wealth level. It is optimal to assign the good to those consumers with low costs. In other words, a consumer is assigned the good if and only if his cost is below a threshold. The wealth information is irrelevant; quantity rationing alone does not allow the public supplier to alter consumers' wealth.

---

<sup>1</sup>The Lagrangean is  $L = m_1[1 - \phi_1(c)] + m_2[1 - \phi_2(c)] + \lambda \{B - m_1[1 - \phi_1(c)]c - m_2[1 - \phi_2(c)]c\}$ , and the first-order derivative with respect to  $\phi_i$  is  $-m_i(1 - \lambda c)$ , which is increasing in  $c$ . Hence, when  $c$  is above a threshold, the first-order derivative is positive, and  $\phi_i = 1$ .

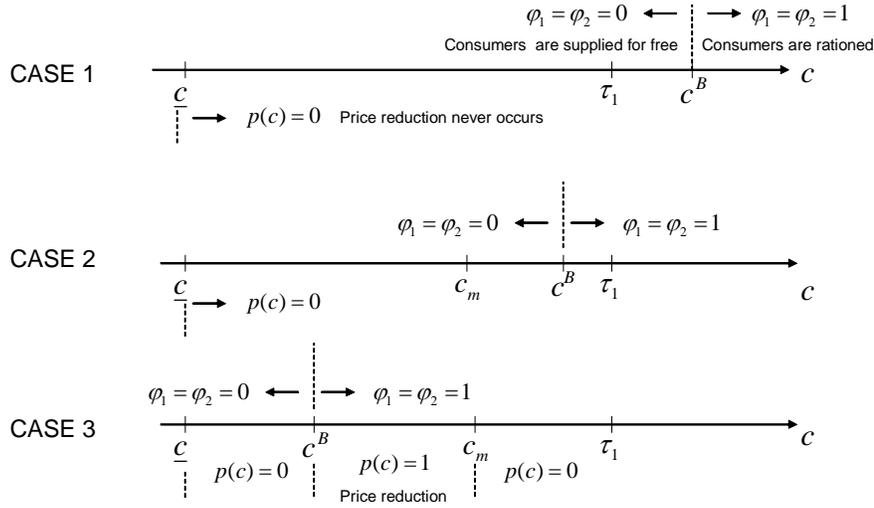


Figure 3: Equilibrium rationing and price reduction

The surprising result in Proposition 3 is that the cost-effectiveness rule continues to hold when the private firm bases its monopoly pricing rule on costs. It is as if the public supplier had ignored the price reactive private firm, and the private firm continued to use its monopoly pricing rule for the customers available. How does this result come about?

Rationing consumers is equivalent to releasing them to the private market. When there are more poor consumers in the private market, the private firm may reduce the price. When there are more rich consumers, the private firm may not. Constraint (16) imposes conditions on the shares of rich and poor consumers for a price reduction to be implemented. Constraint (17) imposes conditions on the share of rich and poor consumers for a price increase to be implemented. Clearly, constraint (17) never binds -raising prices to consumers is not in the public supplier's interest- but constraint (16) may.

There are three factors affecting the public supplier's objective when rationing consumers. First, the cost-effectiveness principle continues to influence rationing policies; the public supplier tends to assign the good to low-cost consumers. Second, rationing poor consumers tends to reduce the price, while rationing rich consumers tends to raise the price. Rationing policies result in the private firm's best responses, which are in (16). Third, rationed rich consumers may gain inframarginal surplus if the price is low. This effect is

absent for poor consumers because the low price in the private market is their willingness to pay.

These three effects can be seen from the first-order derivatives of the Lagrangean  $L$  of Program R with respect to the rationing probabilities  $\phi_1$  and  $\phi_2$  (with the constraint (17) omitted). (The expression of  $L$  is in the proof of the proposition in the appendix.) These derivatives are:

$$\frac{\partial L}{\partial \phi_1} = -m_1(1 - \lambda c) + \mu(c)p(c)m_1[\tau_1 - c] \quad (18)$$

$$\frac{\partial L}{\partial \phi_2} = -m_2(1 - \lambda c) - \mu(c)p(c)m_2[\tau_2 - \tau_1] + m_2p(c)\Delta, \quad (19)$$

where  $\Delta \equiv [U(w_2 - \tau_1) + 1 - U(w_2)] > 0$  is the inframarginal gain for a rich consumer buying at the low price,  $\lambda > 0$  the multiplier for the budget constraint (15), and  $\mu(c) \geq 0$  is the multiplier for the price-reduction constraint (16) at  $c$ .

The common, first term  $-m_i(1 - \lambda c)$  in (18) and (19) is the cost-effectiveness principle. When  $c$  is small, this tends to be negative, so rationing low-cost consumers is unattractive. The third term in (19) captures the inframarginal gain for rich consumers. This is positive if and only if there is a price reduction, when  $p(c) > 0$ . There is not such a corresponding term for the first-order derivative (18) for poor consumers, who never obtain a surplus from the private market.

Each of the second terms in (18) and (19) involves the multiplier  $\mu(c)$  for the price-reduction constraint (16), and the probability of price reduction  $p(c)$ . The interaction between price reduction and inframarginal gain for rich consumers is the key to understanding Proposition 3. First, in (18), this term  $\mu(c)p(c)m_1[\tau_1 - c]$  is positive, and confirms the positive price-reduction effect of rationing poor consumers. By contrast, in (19), the term  $-\mu(c)p(c)m_2[\tau_2 - \tau_1]$  is negative, and confirms the negative price effect of rationing rich consumers.

The critical consideration in Proposition 3 is the conditions under which the public supplier would ration both rich and poor consumers. Consider consumers with a given cost level  $c$ . If all poor consumers with cost  $c$  are supplied, only rich consumers are in the market and the price is  $\tau_2$ , which yields no surplus for rich consumers. If some but not all rich consumers with cost  $c$  are supplied, the inframarginal gain in the private market is relevant. Price reduction and rationing of both rich and poor consumers must occur simultaneously if the inframarginal gain is to be realized. When will this happen?

In Case 1, the budget is large so that it is cost effective to supply some consumers with costs above  $\tau_1$ . Implementing a price reduction at a cost higher than  $\tau_1$  is impossible. Implementing a price reduction at a cost below  $\tau_1$  yields less inframarginal surplus than the gain from the cost effectiveness consideration. Hence, there is never a price reduction under Case 1.

In Case 2, the budget may cover some consumers with costs above  $c_m$ , but there is no price reduction at costs below  $\tau_1$ . For this to happen, we must have (18) positive so that poor consumers are rationed. Also, (19) must also be positive so that rich consumers would be rationed. The Lagrangean  $L$  is linear in  $\phi_i$ , and this implies that all rich and poor consumers are in the market. Nevertheless, at cost  $c > c_m$ , when all rich and poor consumers are in the market, the price-reduction constraint (16) is violated. To implement price reduction requires supplying rich consumers, and this is the opposite of rationing them for the inframarginal surplus. The conflict between the price-reduction and inframarginal effects together means that implementing price reduction is suboptimal.

In Case 3, the budget is very small. According to the cost effectiveness principle, some consumers with costs below  $c_m$  are rationed. At these low costs, the private firm charges  $\tau_1$  when all rich and poor consumers are in the market. There is now no conflict between the price-reduction and inframarginal effects. This is then the only case when price reduction occurs in equilibrium.

Equity concern may lead the public supplier to ration rich consumers more often. Again, we can let the public supplier's payoff be :

$$(1 + \epsilon) m_1 U(w_1) + m_2 U(w_2) + \int_{\underline{c}}^{\bar{c}} \{(1 + \epsilon) m_1 [1 - \phi_1(c)] + m_2 [1 - \phi_2(c)]\} dG(c) \\ + \int_{\underline{c}}^{\bar{c}} m_2 \phi_2(c) p(c) [U(w_2 - \tau_1) + 1 - U(w_2)] dG(c). \quad (20)$$

The optimal policy maximizes (20) subject to (15), (16), (17), and the boundary conditions  $0 \leq \phi_i(c)$ ,  $p(c) \leq 1$ ,  $i = 1, 2$ , each  $c$  in  $[\underline{c}, \bar{c}]$ , and  $p(c) = 0$  for  $c > \tau_1$ . By the Maximum Theorem, the solution is continuous in  $\epsilon$ . Therefore, Proposition 3 is robust.

When the public supplier is concerned with equity, the equilibrium rationing policies favor the poor.

There are two cost thresholds, one for poor consumers, and another for the rich, respectively,  $c^{B1}$  and  $c^{B2}$ , with  $c^{B1} > c^{B2}$ ; consumers are rationed when their costs are higher than the respective levels, so poor consumers are rationed less often than rich consumers.<sup>2</sup>

Figure 4 illustrates a case where some price reduction occurs. Compared to Case 3 in Figure 3, less poor consumers are available in the private market, and the range of costs in which price reduction occurs has shrunk.

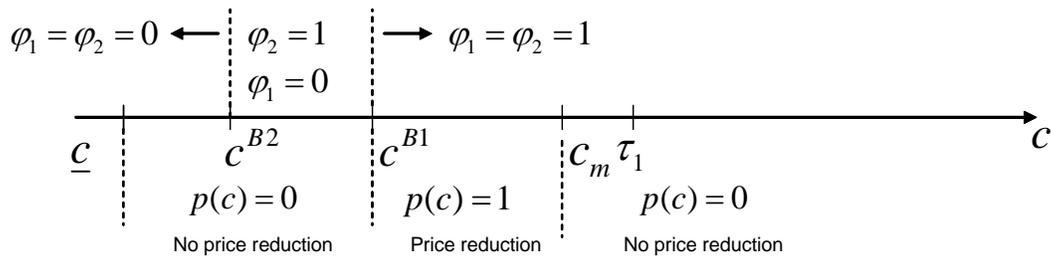


Figure 4: Price reduction with welfare weights.

## 5 Concluding remarks

We have presented a model for studying the effect of rationing on prices in the private market. Public policies should take into account market responses. We have shown that if rationing is based on wealth information, the optimal policy must implement a price reduction in the private market. This is achieved by leaving some poor consumers in the private market. If the public supplier observes consumers' wealth and cost, optimal rationing is based on cost-effectiveness. This framework can be used by the policy maker in the design of public programmes, for instance to anticipate common phenomena such as crowding-out.

The model has been simplified to discrete wealth levels to make the analysis tractable. Extending the model and deriving the equilibrium rationing scheme for a general distribution of consumers' wealth involve more complex computation, but may well be worthwhile. Including quality differences between the public

---

<sup>2</sup>Equilibrium  $c^{B1}$  and  $c^{B2}$  are defined by these two equations:  $c^{B1}/(1 + \epsilon) = c^{B2}$ , and  $\int_{\underline{c}}^{c^{B1}} m_1 c dG(c) + \int_{\underline{c}}^{c^{B2}} m_2 c dG(c) = B$ . The first equation is a cost effectiveness tradeoff adjusted by an equity concern. The second equation is the budget constraint when all poor consumers with cost above  $c^{B1}$  and rich consumers with cost above  $c^{B2}$  are rationed.

and private sector may also be of interest.

# Appendix

**Proof of Proposition 1:** Because all terms in square brackets in the objective function (7) are constant, we can alternatively write the objective function as  $m_2\theta_2G(c_1)$ .

The boundary conditions  $\underline{c} \leq c_1$ , and  $0 \leq \theta_i$  do not bind. If either  $\theta_2 = 0$  or  $c_1 = \underline{c}$  at a solution, then the optimized value is  $m_2\theta_2G(c_1) = 0$ . We show that a rationing policy with  $\theta_1 = \theta_2 = k > 0$  does strictly better. This policy satisfies the budget constraint (8) for some  $0 < k < 1$ . Moreover, from (4) and (5), we have  $c_1 = c_m > \underline{c}$  by assumption. Therefore, this rationing policy,  $\theta_1 = \theta_2 = k$ , is feasible, and yields a payoff  $m_2kG(c_m) > 0$ . This implies that at a solution  $c_1 > \underline{c}$  and  $\theta_2 > 0$ . Because  $c_1 > \underline{c}$ , it follows from (4) that  $\theta_1$  must be bounded away from 0. ■

**Proof of Proposition 2:** From Proposition 1 we know that  $\theta_i > 0$  and  $c_1 > \underline{c}$ . For the time being, ignore the (remaining) boundary conditions  $\theta_i \leq 1$ . Rewrite the budget constraint (8) as  $m_1\theta_1 + m_2\theta_2 \geq m_1 + m_2 - \beta \equiv K > 0$ . Clearly, the budget constraint must bind at a solution. From constraint (4), we have  $m_2\theta_2(\tau_2 - \tau_1) = m_1\theta_1(\tau_1 - c_1)$ , which yields

$$m_1\theta_1 = m_2\theta_2 \frac{\tau_2 - \tau_1}{\tau_1 - c_1}. \quad (21)$$

Substituting this into the modified budget constraint, we can solve for  $m_2\theta_2$ :

$$m_2\theta_2 = K \frac{\tau_1 - c_1}{\tau_2 - c_1}. \quad (22)$$

We next substitute (22) into the objective function  $m_2\theta_2G(c_1)$ . The constrained maximization problem (with the boundary conditions  $\theta_i \leq 1$  omitted) is the same as the unconstrained maximization problem:

$$\max_{c_1} K \frac{\tau_1 - c_1}{\tau_2 - c_1} G(c_1).$$

Ignoring the parameter  $K$ , after simplification we obtain the first-order derivative

$$\frac{g(c_1)}{\tau_2 - c_1} \left[ (\tau_1 - c_1) - \frac{G(c_1)}{g(c_1)} \frac{\tau_2 - \tau_1}{\tau_2 - c_1} \right].$$

Setting the first-order derivative to zero, we have

$$\frac{G(c_1)}{g(c_1)} = \frac{(\tau_1 - c_1)(\tau_2 - c_1)}{\tau_2 - \tau_1}. \quad (23)$$

The left-hand side of (23) is increasing in  $c_1$ . For  $c_1$  between  $\underline{c}$  and  $\tau_1$ , the right-hand side is decreasing. At  $c_1 = \underline{c}$ , the left-hand side of (23) is zero, while the right-hand side of (23) is strictly positive. At  $c_1 = \tau_1$ , the left-hand side of (23) is strictly positive, while the right of (23) is zero. Therefore, there exists a unique  $c_1^*$  strictly between  $\underline{c}$  and  $\tau_1$  that satisfies (23).

To recover  $\theta_i m_i$ , we use (21) and (23) to get  $m_2 \theta_2 = m_1 \theta_1 \frac{G(c_1^*)}{g(c_1^*) (\tau_2 - c_1^*)}$ , which, together with the budget constraint (8), can be used to solve for the values of  $m_1 \theta_1$  and  $m_2 \theta_2$ :

$$m_1 \theta_1 = \frac{m_1 + m_2 - \beta}{1 + \frac{G(c_1^*)}{g(c_1^*)} \frac{1}{\tau_2 - c_1^*}} \quad \text{and} \quad m_2 \theta_2 = \frac{m_1 + m_2 - \beta}{1 + \frac{g(c_1^*)}{G(c_1^*)} (\tau_2 - c_1^*)}. \quad (24)$$

If  $\beta$  is sufficiently large, the right-hand side values in (24) will be less than  $m_1$  and  $m_2$ , and the omitted boundary conditions  $\theta_i \leq 1$  are satisfied. Otherwise, if  $\beta$  is small, one or both of the right-hand side values in (24) will be more than  $m_1$  or  $m_2$ . In this case, a boundary condition binds. ■

**Proof of Lemma 2:** Consider any continuation equilibrium. In this equilibrium, at cost  $c$  the firm will charge either  $\tau_1$  or  $\tau_2$  depending on whether (10) is satisfied. If we have defined  $p$  using the method indicated before the statement of the Lemma, inequalities (11) and (12) are satisfied.

Conversely, let a function  $p : [\underline{c}, \tau_1] \rightarrow [0, 1]$  satisfy inequalities (11) and (12). We show that it characterizes a continuation equilibrium of policy  $(\phi_1, \phi_2)$ . Suppose that  $p(c) = 1$ . Inequality (12) is satisfied by any  $\phi_1(c)$  and  $\phi_2(c)$ . Inequality (11) requires the term inside the curly brackets to be positive, and this means that (10) is satisfied. Next, suppose that  $p(c) = 0$ . Inequality (11) is always satisfied. Inequality (12) requires the term inside the curly brackets in (12) to be negative, and this means that (10) is violated. Last, if  $p(c)$  is a number strictly between 0 and 1, both (11) and (12) must hold as equalities, so that (10) must be an equality. Each value of  $p(c)$  satisfying (11) and (12) corresponds to a continuation equilibrium price. ■

**Proof of Proposition 3:** We use pointwise optimization to solve for the solution of Program R, which is the optimal rationing policy. To do so, we consider a relaxed program in which constraint (17) is omitted; we will show that in the solution of the relaxed program constraint (17) is satisfied. To simplify notation, we multiply (16) by  $g(c)$ , so that  $g(c)$  can be ignored for pointwise optimization. We also write  $\Delta \equiv U(w_2 - \tau_1) + 1 - U(w_2)$ . Let  $\lambda$  denote the multiplier for the budget constraint (15), and  $\mu(c)$  the multiplier

for (16) at  $c$ . The Lagrangean is

$$\begin{aligned} L &= m_1[1 - \phi_1(c)] + m_2[1 - \phi_2(c)] + m_2\phi_2(c)p(c)\Delta \\ &+ \lambda \{B - m_1[1 - \phi_1(c)]c - m_2[1 - \phi_2(c)]c\} \\ &+ \mu(c)p(c) \{m_1\phi_1(c)[\tau_1 - c] - m_2\phi_2(c)[\tau_2 - \tau_1]\}, \end{aligned}$$

where we have omitted the boundary conditions on  $\phi_i$  and  $p$ .

For  $c > \tau_1$ ,  $p(c) = 0$ , so there is no need to optimize over  $p$ , and the first-order derivatives are

$$\frac{\partial L}{\partial \phi_1} = -m_1 + \lambda m_1 c \quad (25)$$

$$\frac{\partial L}{\partial \phi_2} = -m_2 + \lambda m_2 c. \quad (26)$$

For  $c < \tau_1$ , the first-order derivatives are

$$\frac{\partial L}{\partial \phi_1} = -m_1 + \lambda m_1 c + \mu(c)p(c)m_1[\tau_1 - c] \quad (27)$$

$$\frac{\partial L}{\partial \phi_2} = -m_2 + \lambda m_2 c - \mu(c)p(c)m_2[\tau_2 - \tau_1] + m_2 p(c)\Delta \quad (28)$$

$$\frac{\partial L}{\partial p} = m_2\phi_2(c)\Delta + \mu(c) \{m_1\phi_1(c)[\tau_1 - c] - m_2\phi_2(c)[\tau_2 - \tau_1]\}. \quad (29)$$

We consider three cases, according to the size of the budget.

*Case 1* is when the budget is large:  $c^B > \tau_1$ ; that is, the budget is sufficient to cover costs up to a level above poor consumers' willingness to pay  $\tau_1$ . To prove the proposition, we set  $\lambda = \frac{1}{c^B}$ . Now consider  $c > c^B$ . The first-order derivatives (25) and (26) become  $-m_1 + m_1 \frac{c}{c^B}$ , and  $-m_2 + m_2 \frac{c}{c^B}$ , respectively. Both are strictly positive. Hence it is optimal to set  $\phi_i(c) = 1$ . Next, consider  $\tau_1 < c < c^B$ . Then the first-order derivatives (25) and (26) become strictly negative, and it is optimal to set  $\phi_i(c) = 0$ .

Now consider  $\underline{c} < c < \tau_1$ . We claim that  $\phi_i(c) = p(c) = 0$ . At these values, the derivatives (27), (28), and (29) are negative. At  $\phi_i(c) = 0$ , the derivative (29) is zero; hence it is optimal to set  $p(c) = 0$ . At  $p(c) = 0$ , (27) and (28) reduce to  $-m_1 + m_1 \frac{c}{c^B}$ , and  $-m_2 + m_2 \frac{c}{c^B}$ , respectively, and both are strictly negative. It is optimal to set  $\phi_i(c) = 0$ . Finally, the omitted constraint (17) is satisfied since  $\phi_i(c) = 0$ .

*Case 2* is when the budget  $c^B$  is lower, between  $c_m$  and  $\tau_1$ :  $c_m < c^B < \tau_1$ . Recall that  $c_m$  is the cost level at which the firm will set the low price  $\tau_1$  if it has access to all consumers ( $m_1[\tau_1 - c_m] = m_2[\tau_2 - \tau_1]$ ),

see also (5)). Again we set  $\lambda = \frac{1}{c^B}$ . For  $c > \tau_1$ , the first-order derivatives (25) and (26) are  $-m_1 + m_1 \frac{c}{c^B}$ , and  $-m_2 + m_2 \frac{c}{c^B}$ , respectively. Both are strictly positive. Hence it is optimal to set  $\phi_i(c) = 1$ .

Next, consider  $c^B < c < \tau_1$ . We set  $\mu(c)$  to satisfy

$$m_2 \Delta + \mu(c) \{m_1[\tau_1 - c] - m_2[\tau_2 - \tau_1]\} = 0. \quad (30)$$

Because  $c > c^B > c_m$ , we have  $m_1[\tau_1 - c] < m_2[\tau_2 - \tau_1]$ . Therefore,  $\mu(c) > 0$ . We claim that  $p(c) = 0$ ,  $\phi_i(c) = 1$ . Given  $p(c) = 0$ , the first-order derivatives (27) and (28) are  $-m_1 + m_1 \frac{c}{c^B}$ , and  $-m_2 + m_2 \frac{c}{c^B}$ , respectively. Both are strictly positive. Hence it is optimal to set  $\phi_i(c) = 1$ . Given  $\phi_i(c) = 1$ , by the choice of  $\mu(c)$  satisfying (30), the derivative (29) is zero. Hence, setting  $p(c) = 0$  is optimal. Obviously, the omitted constraint (17) is satisfied since  $\phi_i(c) = 1$ .

Next, consider  $\underline{c} < c < c^B$ . We claim that  $\phi_i(c) = p(c) = 0$ . Given  $p(c) = 0$ , the first-order derivatives (27) and (28) are both negative when  $c < c^B$ . Hence it is optimal to set  $\phi_i(c) = 0$ . Next, given that  $\phi_i(c) = 0$ , the derivative (29) is zero. Hence it is optimal to set  $p(c) = 0$ . Again, the omitted constraint (17) is satisfied since  $\phi_i(c) = 0$ .

*Case 3* is when the budget is small,  $c^B < c_m$ . We set  $\lambda = \frac{1}{c^B}$ . For  $c > \tau_1$ , we use the same argument as in Case 1 and Case 2, and  $\phi_i(c) = 1$ . For  $c_m < c < \tau_1$ , we claim that  $\phi_i(c) = 1$  and  $p(c) = 0$ . We show this using the same argument as in Case 2. When  $\mu(c)$  is set to be sufficiently large, the first-order derivative (29) is zero, so that  $p(c) = 0$  is optimal when  $\phi_i(c) = 1$ . When  $p(c) = 0$ , setting  $\phi_i(c) = 1$  is optimal. The omitted constraint (17) is satisfied because  $\phi_i(c) = 1$  and  $c > c_m$ .

Next, for  $c^B < c < c_m$ , we claim that  $p(c) = 1$  and  $\phi_i(c) = 1$ . We set  $\mu(c) = 0$ . When  $\phi_i(c) = 1$ , the first-order derivative (29) becomes

$$\frac{\partial L}{\partial p} = m_2 \Delta > 0,$$

and it is optimal to set  $p(c) = 1$ . Given  $p(c) = 1$  and  $\mu(c) = 0$ , first-order derivatives (27) and (28) are strictly positive since  $c^B < c$ . Hence, it is optimal to set  $\phi_i(c) = 1$ . The omitted constraint (17) is satisfied because  $p(c) = 1$ .

Finally, for  $\underline{c} < c < c^B$ , we claim that  $\phi_i(c) = p(c) = 0$ . Given  $p(c) = 0$ , the first-order derivatives (27)

and (28) are strictly negative because  $c < c^B$ . Hence it is optimal to set  $\phi_i(c) = 0$ . Given  $\phi_i(c) = 0$ , the first-order derivative (29) is zero. It is optimal to set  $p(c) = 0$ . The omitted constraint (17) is satisfied because  $\phi_i(c) = 0$ . ■

## References

Barros, Pedro Pita, and Pau Olivella, "Waiting Lists and Patient Selection," *Journal of Economics & Management Strategy*, 14(3), (2005), 623-646.

Beato, Paulina and Andreu Mas-Colell , "The Marginal Cost Pricing as a Regulation Mechanism in Mixed Markets." In: M. Marchand, P. Pestieau and H. Tulkens, Editors, *The performance of public enterprises*, North-Holland, Amsterdam (1984), 81–100.

Cremer, Helmuth, Marchand, Maurice and Thisse, Jacques-Francois, "Mixed Oligopoly with Differentiated Products", *International Journal of Industrial Organization*, 9(1), (1991), 43-53.

Cutler, David M. and Jonathan Gruber, "Does Public Insurance Crowd Out Private Insurance?", *Quarterly Journal of Economics*, 111(2), (1996), 391-430

Defraja, Gianni and Flavio Delbono, "Game Theoretic Models of Mixed Oligopoly", *Journal of Economic Surveys*, 4(1), (1990), 1-17.

Grassi, Simona and Ching-to Albert Ma, "Public Rationing and Private Sector Selection", Boston University working paper, (2008).

Gruber, Jonathan and Kosali Simon, "Crowd-out 10 Years Later: Have Recent Public Insurance Expansions Crowded Out Private Health Insurance?", *Journal of Health Economics*, 27, (2008), 201-217.

Hoel, Michael, "What Should (Public) Health Insurance Cover?," *Journal of Health Economics*, 26, (2007), 251-262.

Hoel, Michael, and Erik Magnus Sæther, "Public Health Care with Waiting Time: the Role of Supplementary Private Health Care," *Journal of Health Economics*, 22, (2003), 599-616.

Iversen, Tor, "The Effect of a Private Sector on the Waiting Time in a National Health Service," *Journal of Health Economics*, 16, (1997), 381-396.

Lo Sasso, Anthony and Bruce D. Meyer, "The Health Care Safety Net and Crowd-Out of Private Health Insurance", NBER Working Papers, 11977, (2006).

Merrill, William C. and Norman Schneider, "Government Firms in Oligopoly Industries: A Short-Run Analysis", *Quarterly Journal of Economics*, 80, (1966), 400–12.

Rask, Kevin N. and Kimberly J. Rask, "Public Insurance Substituting for Private Insurance: New Evidence Regarding Public Hospitals, Uncompensated Care Funds, and Medicaid", *Journal of Health Economics*, 19, (2000), 1-31.

Shore-Sheppard, Lara D., "Stemming the Tide? The Effect of Expanding Medicaid Eligibility on Health Insurance", *The B.E. Journal of Economic Analysis & Policy*, Advances, (2009), 8.2.