Optimal Public Rationing and Price Response

Simona Grassi

Max Weber Fellow
European University Institute
Villa La Fonte, Via Delle Fontanelle, 10
I - 50014 San Domenico di Fiesole (FI)
Italy

Simona.Grassi@eui.eu

Ching-to Albert Ma

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

ma@bu.edu

August 2009

Abstract

We study optimal public rationing of an indivisible good and private sector price responses. Consumers differ in their wealth and costs 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 consumers' 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 consumers' wealth and cost information. In equilibrium, consumers are allocated the good if and only if their costs are below a threshold. Wealth information is not used. Rationing based on cost results in higher equilibrium total consumer surplus than rationing based on wealth.

1 Introduction

Public supply of services and goods such as health, education, and housing is very common. Because of limited budgets, these goods cannot be freely provided to all consumers and some rationing must occur. Rationed consumers may turn to the private market and purchase at their own expense. In this paper we study the optimal rationing policy and price responses in the private market.

The design of an optimal rationing policy should take into account private market reactions to a public program. For example, in the public finance literature, a major policy concern is "crowd out," in which an increase in public supply may decrease consumers' purchases in the private market, and may result in a lower total quantity consumed. Following Cutler and Gruber (1996), many studies have measured the extent to which an expansion in Medicaid and similar programs for the indigent may crowd out private insurance coverage. In the last twenty years in the US, both the numbers of uninsured and publicly insured individuals have risen (Gruber and Simon, 2008). Although estimates of health insurance crowd out vary, most empirical studies agree that an increase in the publicly insured population is associated with a decrease in the privately insured. In this paper we exhibit a mechanism in which price increases in the private sector may be best responses against public supply. We can explain a reduction in private market quantity when the public supply expands.

In our model, consumers are heterogenous in two dimensions: they have different wealth levels, and the costs of providing the good to them differ. Wealth heterogeneity is a natural assumption. For our model, wealth heterogeneity means that rich consumers are more willing to pay for services than poor consumers. Provision cost heterogeneity is also common. In the health market, 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 normalize by stipulating that one unit of the good represents one unit of improvement in utility. We assume that the utility benefit is the same to rich and poor consumers.

Some examples illustrate these assumptions. Consider a hip replacement procedure. Treating patients with higher levels of disability requires higher costs. A hip replacement will allow a consumer to walk about without pain, which we define as a unit of health improvement. Consumers who are more wealthy may be more willing to pay for the hip replacement. Now consider education to achieve literacy. Different students will require different costs for this goal, and rich families are more willing to pay for education to achieve it. Literacy can be thought of as the good to be provided to students, and we normalize it so that it represents a unit of benefit.

We derive optimal rationing policies and price reactions under two information regimes. In the first, the public supplier observes consumers' wealth level and can credibly commit to a rationing scheme based on wealth. In the second, the public supplier observes both consumers' wealth and cost levels and credibly commits to a rationing scheme based on both pieces of information. In each regime, we study the following extensive form. First, the public supplier chooses a rationing scheme. Second, the private firm, unable to observe consumers' wealth levels, sets its prices according to consumers' cost of provision. Third, consumers who are rationed by the public supplier may purchase from the private firm. We use aggregate consumer utility as the public supplier's preferences, and assume that the private market consists of a profit-maximizing monopolist.

Rationing determines who among consumers are entitled to public provision. In the first regime, in equilibrium, when rationing is based on wealth, the public supplier must ration both poor and rich consumers, and implement price reduction in the private sector. What is the intuition behind this result? If poor consumers are supplied and rich consumers are rationed, then only rich consumers will be in the private market. The private firm cream-skims consumers by setting a high price. The public supplier can mitigate cream-skimming by rationing some poor consumers, making them available to the private market. The firm may then find it attractive to set a low price when costs are low. In other words, to limit price increase, the public supplier optimally chooses to leave a mixture of rich and poor consumers in the private market.

Those concerned with equity may find rationing the poor disagreeable. Nevertheless, rationing some poor consumers always yields a first-order gain due to price reduction. Spending the budgets only on poor consumers will exacerbate cream-skimming, eliminating this gain. We can extend the analysis to consider the optimal policy of a public supplier who values the utility of the poor more than the rich. Price reduction then is implemented less often, but unless the social preferences put very high weights on poor consumers, the price reduction effect from rationing *some* poor consumers remains beneficial.

In the second information regime, rationing can be based on both wealth and cost information. Clearly, the public supplier's equilibrium payoff must be higher compared to rationing based only on wealth. Surprisingly, in equilibrium the public supplier rations consumers according to cost information alone, ignoring wealth information altogether. We derive a *cost effectiveness* principle, which aims to get the "biggest bang for the buck:" all consumers with costs below a threshold are supplied while those with higher costs are rationed. By assumption, a unit of the (indivisible) good confers one unit of utility to a consumer, irrespective of wealth status. The most efficient use of the public budget is to serve the largest number of consumers. This is achieved by spending the budget on consumers with lower costs. Using rationing to implement price reduction is suboptimal because cost effectiveness is already achieved.

The private market is an option for higher-cost consumers who are willing to pay for the good, and remains so even if it sets a high price. Under the equilibrium rationing scheme, the wealth distribution of consumers in the market is the *ex ante* distribution. The firm cannot do better than setting prices as if the public sector did not exist, although it only sells to higher-cost consumers.

Clearly, if the public supplier can only pick one piece of information for rationing, it will choose cost rather than wealth information. Once cost information is available, wealth information does not improve the design of optimal rationing. Crowd out—higher prices in the private sector—is less of a concern when the public supply can be based on costs.

In equilibrium, poor and rich consumers are treated equally because public supply is only based on costs, and the public supplier's preferences are aggregate consumer utility. Again, we can let the public supplier value poor consumers' utility more than rich consumers'. In equilibrium, the cost effectiveness principle will be slightly different: the rationing cost threshold for poor consumers will become higher than rich consumers. More poor consumers than rich will be supplied. Price reduction will be implemented even less often.

Our model is unlike the typical regulation model. First, the public supplier does not seek to regulate the private market. Regulating an entire market is a much more complex issue, especially in the health care and education sectors. We focus on a different problem: the interaction between the public supplier's limited budget, the information that is used to supply consumers, and the private firm's price responses against rationing policies. Second, the public supplier may be informed. In the first regime, the public supplier's rationing is based on wealth information, while the private firm sets prices according to costs. In the second, the public supplier has cost information, too. The public supplier is never less informed than the private market. Instead, we study how the public supplier should use a limited budget given two information regimes and a price-reactive private sector.

Our information assumptions are also plausible. The public supplier has access to wealth information through tax returns. It may well have access and use cost information. In our first regime, the public supplier is unable to use cost information. 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. Often free supplies are based on income or wealth only, not on provision costs. Medicaid and other programs target the indigent, and education subsidies are often only available to low-income households. By contrast, we also study optimal rationing based on wealth and cost in the second regime. This allows us to compare the difference between wealth and cost information on optimal rationing design.

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 the public supplier is assumed unable to commit to rationing policies. The public supplier and private firm react to each other, and rationing and price schemes must be mutual best responses. If rationing is based on wealth, the game in Grassi and Ma (2008) has a continuum of 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. Also, we find that when the public supplier observes consumers' wealth and cost levels, it implements a rationing policy based on cost effectiveness. Under rationing based on wealth and cost information, equilibria are robust against changes in the public supplier's commitment power.

The literature on rationing and its interaction with the private market assumes an exogenous supply in the private market. Barros and Olivella (2005) consider doctors working in the public sector self-refer patients to their private practices. 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. The common assumption has been a competitive private market. This is unlikely to be true in most settings. 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 how public policies influence private market responses.

Our model provides a mechanism for the crowd out of private purchases due to the expansion of free public provision in the health sector. Cutler and Gruber (1996) informally suggest some "mechanisms through which employer-provided coverage could fall as Medicaid eligibility increases." Although they suggest that the insurance premium may increase in the private market, the precise mechanism for this price increase has not been spelled out. Our model provides such a mechanism.

Strategic interaction between public and private sectors has been studied in the literature of mixed oligopolies. In a mixed oligopoly 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 MasColell (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 prices below costs. This literature has focused on public utilities, such as telecommunication, transportation, water, and energy where mixed oligopolies are common. A mixed oligopoly is also common in health care, education and housing sectors. Issues of rationing and cream-skimming are especially relevant in the health care sector, but have not been the focus of that literature. By contrast, we consider rationing and different information structures.

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 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 nothing), 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 τ_i is defined by the following

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

so τ_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, $G : [\underline{c}, \overline{c}] \to [0,1]$ its distribution function, and g the corresponding density, both defined on a positive support, and $g > 0$. The distribution G is independent of wealth and uncorrelated to the utility benefit from consuming the good (which is fixed at 1). Furthermore, we assume that the hazard rate $G(c)/g(c)$ is increasing. Let γ be the expected value of c. We assume that $\underline{c} < \tau_1 < \tau_2 = \overline{c}$; the equality $\tau_2 = \overline{c}$ involves no loss of generality because a firm will never be able to sell to any consumer when cost is above τ_2.

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 nonprice rationing mechanism based on wealth. In the second, the public supplier uses a nonprice rationing mechanism based on both wealth and cost.

When rationing is based on consumers' wealth, a rationing policy is given by (θ_1, θ_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 (ϕ_1, ϕ_2), a pair of functions $\phi_i : [\underline{c}, \overline{c}] \to [0, 1]$, $i = 1, 2$. 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

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

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

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

Consumers who are rationed may consider buying from the private market.

The public supplier's payoff is the sum of consumer utilities. Again, we focus on the optimal public supply, not the optimal regulation of the entire market. Therefore, it is natural to assume that the public supplier is concerned with consumer surplus. We will mostly be concerned with a simple sum of consumer surplus, but will discuss how our results will change when the utility of poor consumers is given a higher weight than the utility of rich consumers.

There is a private market which we model as a monopoly for convenience. The firm observes a consumer's cost c, but not his wealth w_i. In our model, if the firm managed to observe 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 (θ_1, θ_2), $0 \leq \theta_i \leq 1$, supplying $(1-\theta_i)m_i$ of consumers with wealth w_i, $i = 1, 2$. Under rationing based on wealth and cost, the public supplier sets a rationing policy (ϕ_1, ϕ_2), $\phi_i : [\underline{c}, \overline{c}] \to [0, 1]$, supplying $[1 - \phi_i(c)]m_i$ of 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 prices set at Stage 2.

3 Equilibrium rationing and prices in wealth-based rationing

3.1 Continuation equilibrium prices

For a given rationing policy (θ_1, θ_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 at either τ_1 or τ_2. Clearly at any cost above τ_1, the firm must set the price at τ_2, selling only to rich consumers. Suppose that the cost c decreases below τ_1. The firm may set a low price τ_1, selling to both rich and poor consumers, or a high price τ_2, selling only to rich consumers; these profits are respectively:

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

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

The profits in (2) and (3) are linear in c, and the profit in (2) decreases in c at a faster rate than (3); these functions are illustrated in Figure 1. For some rationing policies (θ_1, θ_2) there may be a cost level $c_1 < \tau_1$ in $[\underline{c}, \overline{c}]$ such that $\pi(\tau_1; c_1) = \pi(\tau_2; c_1)$, as in Figure 1. This value of c_1 is given by

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

As the cost drops below τ_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 τ_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 (θ_1, θ_2), in a continuation equilibrium if c_1 in (4) is greater than $\underline{c}$, the firm sets the high price τ_2 if $c > c_1$, and the low price τ_1 if $c < c_1$; if c_1 in (4) is less than $\underline{c}$, the firm always sets the high price τ_2.*

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 $[\underline{c}, \overline{c}]$ where

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

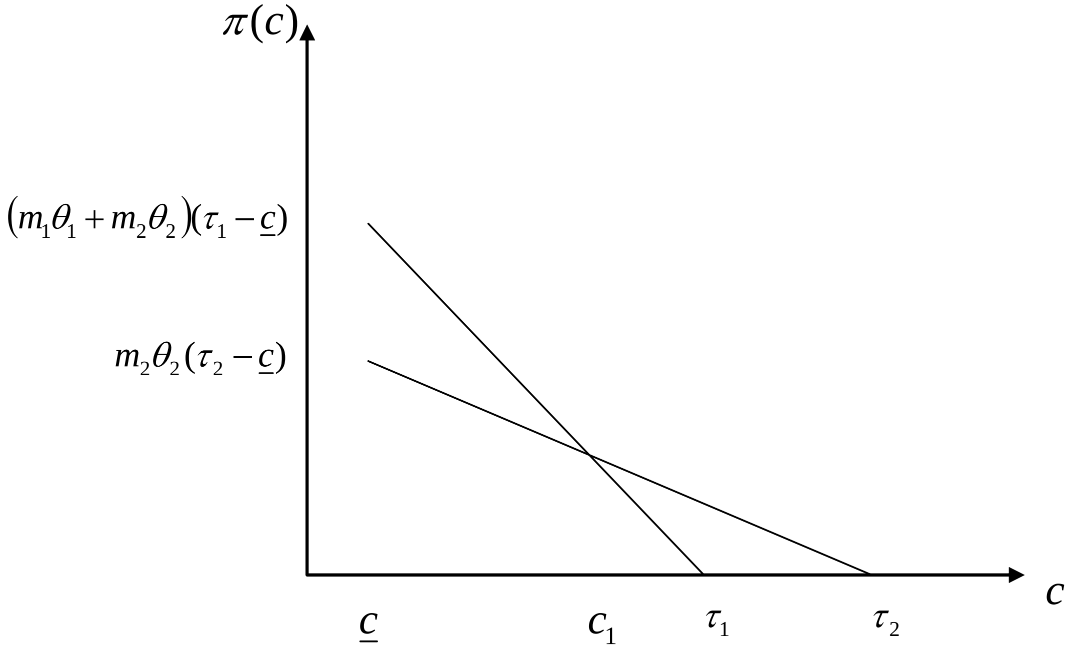

Figure 1: Comparison of profit between setting high and low prices

3.2 Equilibrium rationing

Given the continuation equilibrium prices, the aggregate consumer utility is:

$$m_1\left[(1-\theta_1)[U(w_1)+1]+\theta_1\left\{\int_{\underline{c}}^{c_1}[U(w_1-\tau_1)+1]\ \mathrm{d}G+\int_{c_1}^{\overline{c}}U(w_1)\ \mathrm{d}G\right\}\right]+$$
$$m_2\left[(1-\theta_2)[U(w_2)+1]+\theta_2\left\{\int_{\underline{c}}^{c_1}[U(w_2-\tau_1)+1]\ \mathrm{d}G+\int_{c_1}^{\overline{c}}[U(w_2-\tau_2)+1]\ \mathrm{d}G\right\}\right].$$

In this expression, terms involving $(1-\theta_i)$ are consumers' utilities when they receive the public supply at no charge. Terms involving θ_i are the market outcomes. For poor consumers, if their costs are below c_1, they purchase at τ_1, which actually leaves them no surplus (see definition of τ_i in (1)). Similarly, for rich consumers, if their costs are above c_1, they purchase at price τ_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 τ_1 is lower than their willingness to pay τ_2.

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

$$[m_1U(w_1)+m_2U(w_2)+m_1(1-\theta_1)+m_2(1-\theta_2)]+m_2\theta_2\int_{\underline{c}}^{c_1}[U(w_2-\tau_1)+1-U(w_2)]\,g(c)\mathrm{d}c, \tag{6}$$

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 τ_1.

We introduce a new notation $\beta \equiv B/\gamma$. Because B denotes the budget available to the public supplier and γ the expected cost, the term β is the number of consumers to whom the public supplier can provide the good. In an equilibrium the budget B must be exhausted. Hence, we replace $m_1(1-\theta_1)+m_2(1-\theta_2)$ by β, and simplify (6) to

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

An equilibrium is a rationing policy (θ_1, θ_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

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

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

Proposition 1 *In equilibrium, the public supplier rations consumers in each wealth class: $\theta_1 > 0$ and $\theta_2 > 0$, while the firm charges the low price τ_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 price. Therefore, the equilibrium cost threshold c_1, at which the price drops from τ_2 to τ_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 (θ_1, θ_2).

Proposition 2 *If the budget B is sufficiently large, the value of c_1^* is 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 rations some consumers in each wealth class. If the budget is small, either θ_1 or θ_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. An "interior" solution results when there is enough budget, the value of c_1^* is obtained by a first-order condition, and the boundary conditions $\theta_i \leq 1$ do not bind. When the budget is small, there is a "corner" solution, and either $\theta_1 = 1$ or $\theta_2 = 1$. In the interior solution, the cost threshold at which price reduction occurs remains constant even as the budget increases: c_1^* is independent of the budget.

According to Proposition 2, the equilibrium cost threshold c_1^* may be higher or lower than c_m, the threshold at which price reduction occurs when there is no public supply. Nevertheless, if $\theta_2 = 1$, we must have $\theta_1 < 1$, and $c_1^* < c_m$. Rationing all rich consumers means that the budget must be spent on poor consumers. With less poor consumers in the market, price reduction is less often. With price at the high level for more cost realizations, public supply reduces transactions in the private market. This offers an explanation of crowd out.

We now explain the intuitions behind Proposition 2. The maximization of the objective function in (7) is equivalent to the maximization of $m_2\theta_2 [U(w_2 - \tau_1) + 1 - U(w_2)] G(c_1)$, which is the surplus earned by the fraction θ_2 of rationed rich consumers with costs below the threshold c_1. The inframarginal gain is constant, and the optimization problem is equivalent to the maximization of $m_2\theta_2 G(c_1)$. This objective function is increasing in both θ_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 θ_2 and c_1 can be.

The constraints are graphed in Figure 2. The downward sloping line is the budget constraint (8); it shifts inward as the budget (β) increases. The upward sloping line through the origin is the locus of $m_1\theta_1$ and $m_2\theta_2$ combinations that implement a price reduction at cost threshold c_1; see (4). As the value of c_1 increases, the line becomes flatter. The two dotted lines are the boundary conditions for $m_i\theta_i$.

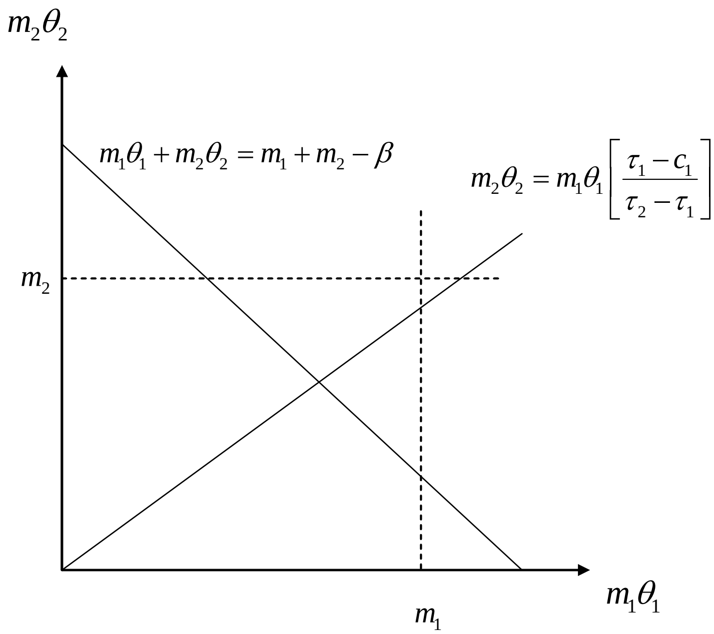

Figure 2: Budget and cost threshold constraints; boundary conditions.

By solving for θ_1 and θ_2 with (4) and (8), and then substituting them 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 θ_i after setting c_1 to c_1^*, and both θ_1 and θ_2 are smaller than 1:

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

(Details of the computation are in the proof.) 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), the total number of consumers m_i in a wealth class determines 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 θ_1 becomes 1. That is, to implement a cost reduction, the public supply 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 θ_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 need to 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 ϵ so our results, Propositions 1 and 2, are robust: a small increase of ϵ 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, and generates a larger extent of crowd out.

We conclude this section by describing the optimal rationing policy 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) where consumers have 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 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

We begin with the continuation equilibria given a rationing policy (ϕ_1, ϕ_2), $\phi_i : [\underline{c}, \overline{c}] \to [0, 1]$. Again, there are only two possible equilibrium prices in the private market, the low price τ_1 and the high price τ_2. For any $c > \tau_1$, the unique best response by the private firm is τ_2. For any c between $\underline{c}$ and τ_1, the firm chooses between the low price τ_1 and the high price τ_2. The firm's profit from the low price τ_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 τ_2. Therefore, the firm sets the low price τ_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

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

It sets a high price τ_2 if (10) is violated, and it may randomize between τ_1 and τ_2 if (10) holds as an equality. These are the continuation equilibrium prices.

We now define an indicator function for continuation equilibria when $c < \tau_1$. Let $p : [\underline{c}, \tau_1] \to [0,1]$. Given a policy (ϕ_1, ϕ_2), we set $p(c) = 1$ if (10) holds as a strict inequality, $p(c) = 0$ 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 cost is between c and τ_1. If $p(c)$ takes the value 0, we understand it to mean no price reduction, and the private firm chooses the high price τ_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 τ_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 τ_1, any continuation equilibrium is given by a function $p : [\underline{c}, \tau_1] \to [0,1]$ satisfying the following two inequalities:*

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

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

Lemma 2 (whose proof is in the Appendix) uses a "complementary" function $p(c)$ to describe the continuation equilibria. The key characterization of continuation equilibria is whether at cost c a price reduction will be implemented by the policy (ϕ_1, ϕ_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 $p : [\underline{c}, \tau_1] \to [0,1]$ satisfying (11) and (12).

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

4.2 Equilibrium rationing

Given a rationing policy (ϕ_1, ϕ_2), $\phi_i(c) : [\underline{c}, \overline{c}] \to [0, 1]$, the continuation equilibrium is characterized by the function p in Lemma 2. For values of c higher than τ_1, the equilibrium price in the private sector must be τ_2, and $p(c) = 0$. For values of c less than τ_1, rationed consumers may purchase at the private sector at the low price when $p(c) = 1$ or at the high price when $p(c) = 0$. We now write down the public supplier's payoff in the continuation equilibrium:

$$\int_{\underline{c}}^{\overline{c}} \{m_1[1-\phi_1(c)][U(w_1)+1] + m_2[1-\phi_2(c)][U(w_2)+1]\} \, \mathrm{d}G(c)$$
$$+ \int_{\underline{c}}^{\overline{c}} m_1\phi_1(c) \{[1-p(c)]U(w_1) + p(c)[U(w_1-\tau_1)+1]\} \, \mathrm{d}G(c)$$
$$+ \int_{\underline{c}}^{\overline{c}} m_2\phi_2(c) \{[1-p(c)][U(w_2-\tau_2)+1] + p(c)[U(w_2-\tau_1)+1]\} \, \mathrm{d}G(c).$$

In this expression, the first integral is the sum of 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 down the sum of utilities of rationed poor consumers. A poor consumer who has cost c will encounter a price reduction with probability $p(c)$ (in a continuation equilibrium). If there is no 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 a price τ_1, hence the term $U(w_1 - \tau_1) + 1$. In the last integral, we write down the sum of utilities of rationed rich consumers. If there is no price reduction, the rich consumer buys at τ_2, hence the term $U(w_2 - \tau_2) + 1$. If there is a price reduction, he buys at τ_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 τ_1. Poor consumers either do not buy or buy at their reservation price τ_1, gaining no surplus from the private market. We use the definitions of τ_1 and τ_2 to simplify the payoff to:

$$m_1U(w_1) + m_2U(w_2) + \int_{\underline{c}}^{\overline{c}} \{m_1[1-\phi_1(c)] + m_2[1-\phi_2(c)]\} \, \mathrm{d}G(c)$$
$$+ \int_{\underline{c}}^{\overline{c}} m_2\phi_2(c)p(c) \, [U(w_2-\tau_1) + 1 - U(w_2)] \, \mathrm{d}G(c). \tag{13}$$

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 τ_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 (ϕ_1, ϕ_2) and the corresponding price reduction function p to maximize

$$\int_{\underline{c}}^{\overline{c}} \{m_1[1-\phi_1(c)] + m_2[1-\phi_2(c)]\} \, \mathrm{d}G(c) + \int_{\underline{c}}^{\overline{c}} m_2\phi_2(c)p(c) \left[U(w_2-\tau_1) + 1 - U(w_2)\right] \mathrm{d}G(c) \quad (14)$$

subject to

$$B - \int_{\underline{c}}^{\overline{c}} \{m_1[1-\phi_1(c)] + m_2[1-\phi_2(c)]\} \, c\mathrm{d}G(c) \geq 0 \quad (15)$$

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

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

and the boundary conditions $0 \leq \phi_i(c) \leq 1$, $i = 1, 2$, $0 \leq p(c) \leq 1$, each c in $[\underline{c}, \overline{c}]$, and $p(c) = 0$ for $c > \tau_1$. 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). Call this Program R.

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,*

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

where the cost threshold c^B is defined by

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

The proposition (whose proof is in the Appendix) says that optimal rationing is based only on cost. Rich and poor consumers are treated equally. Public supply is determined according to a cost effectiveness criterion. Given that the good provides a unit utility increment, in equilibrium the public supplier provides the good for free if and only if the cost-benefit ratio is below a threshold. The cost level c^B in the proposition refers to one at which supplying the good for free to consumers with costs below c^B, irrespective of wealth status, will exhaust the budget.

Before we present the intuition behind Proposition 3, we write down a benchmark where the private market is inactive. Here, the public supplier chooses a rationing policy (ϕ_1, ϕ_2) to maximize consumer's utility subject to the budget constraint. This actually is a special case of this model: simply set $p(c) = 0$ for all c, 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.[1] 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.

The surprising result in Proposition 3 is that the cost effectiveness rule continues to hold when the private firm is a monopolist. 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 on available consumers. 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 price. When there are more rich consumers, the private firm may not. Constraint (16) imposes conditions on the masses of rationed rich and poor consumers to implement a price reduction. Constraint (17) imposes conditions on the masses of rationed rich and poor consumers to implement a price increase. Clearly, constraint (17) never binds—raising prices 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,

[1] 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 ϕ_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 $\phi_i = 1$.

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 price, while rationing rich consumers tends to raise 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 ϕ_1 and ϕ_2 (with the constraint (17) omitted). (The expression for L is in the proof of the proposition in the Appendix.) These derivatives are:

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

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

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 no such corresponding term for the first-order derivative (18) for poor consumers, who never obtain a surplus from the private market.

Each of the second term 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

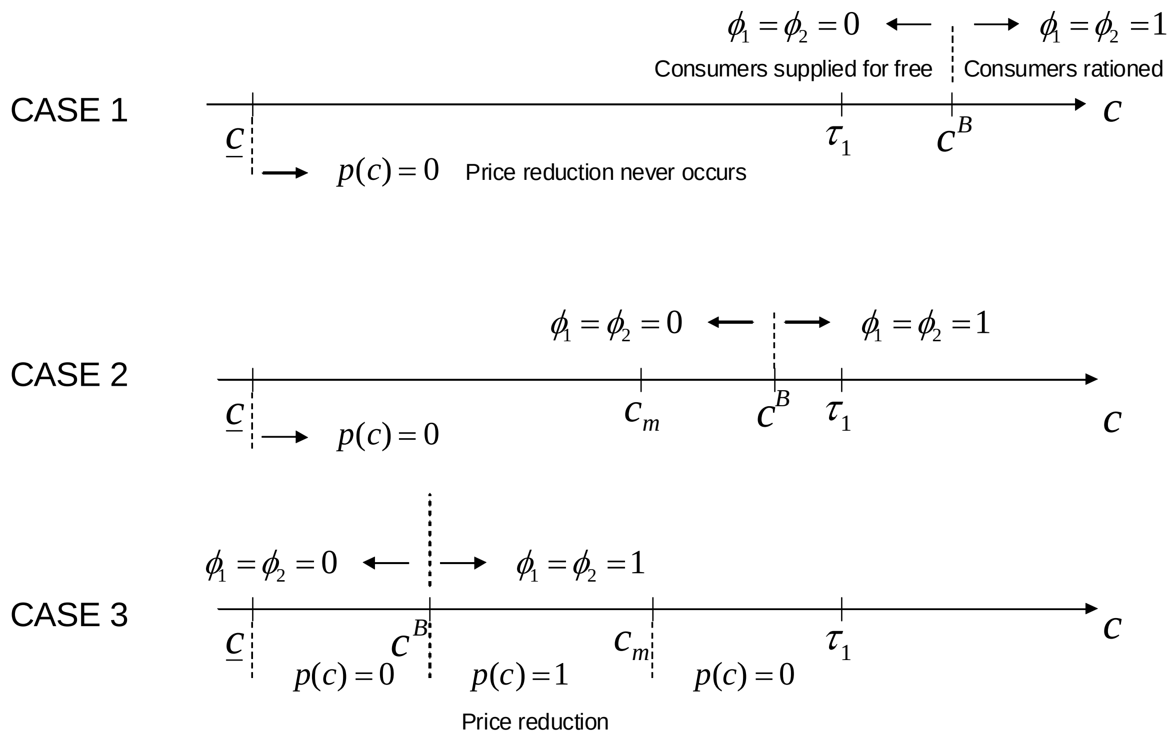

Figure 3: Equilibrium rationing and price reduction

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 τ_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?

Figure 3 shows the three cases that make up the proof of Proposition 3. In Case 1, the budget is large so that it is cost effective to supply some consumers with cost above τ_1. Implementing a price reduction at a cost higher than τ_1 is impossible. Implementing a price reduction at a cost below τ_1 yields less inframarginal surplus than the gain from cost effectiveness consideration. Hence, there is never a price reduction.

In Case 2, the budget is medium sized, and may cover some consumers with cost above c_m, but there is no price reduction at cost below τ_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 ϕ_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 some rich consumers, eliminating the inframarginal gain for these rich consumers in the private market. The conflict between the price-reduction and inframarginal-gain effects together imply that implementing price reduction is suboptimal.

In Case 3, the budget is small. According to the cost effectiveness principle, some consumers with costs below c_m are rationed. At these low costs, the private firm charges τ_1 when all rich and poor consumers are in the market. There is now no conflict between the price-reduction and inframarginal-gain effects. This is then the only case when price reduction occurs in equilibrium.

Clearly, the public supplier's equilibrium payoff—aggregate consumer utility—under rationing based on cost and wealth cannot be lower than rationing based on wealth alone. Proposition 3 establishes that aggregate consumer utility is strictly higher, and that the optimal rationing rule is based only on cost. Once cost information is available, wealth information does not improve the public supplier's payoff. We summarize by the following:

Corollary 1 *Equilibrium aggregate consumer utility is higher under rationing based on cost than wealth. If the public supplier must pick between cost and wealth information to administer rationing, it optimally will choose cost information.*

Once cost information becomes available, price increase is no longer a dominant issue. In equilibrium, the public supplier does not deliberately ration poor consumers in order to implement more price reduction, in order to avoid crowd out. The private sector now is an option for consumers whose costs are too high (relative to the budget), and this does not conflict with the public supplier's objective.

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

$$(1+\epsilon)\, m_1 U(w_1) + m_2 U(w_2) + \int_{\underline{c}}^{\overline{c}} \{(1+\epsilon)\, m_1[1-\phi_1(c)] + m_2[1-\phi_2(c)]\}\, \mathrm{d}G(c)$$
$$+\int_{\underline{c}}^{\overline{c}} m_2\phi_2(c)p(c)\,[U(w_2-\tau_1)+1-U(w_2)]\,\mathrm{d}G(c), \tag{20}$$

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

When the public supplier is concerned with equity, the equilibrium rationing policy favors the poor. There are two cost thresholds, one for the poor, and another for the rich, respectively, c^{B1} and c^{B2}, with $c^{B1} > c^{B2}$. Consumers in each wealth class are rationed when their costs are higher than the respective levels, so poor consumers are rationed less often than rich consumers.[2]

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. Here, the price in the private market is nonmontone in cost. The price is low if and only if cost is between c_m and c^{B1}.

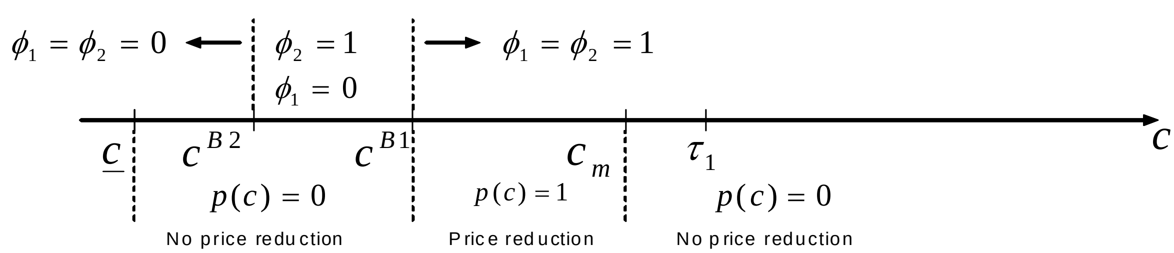

Figure 4: Price reduction with welfare weights.

Finally, we discuss the optimal rationing policy when the private market is competitive so that

[2]Equilibrium c^{B1} and c^{B2} are defined by two equations: $c^{B1}/(1+\epsilon) = c^{B2}$, and $\int_{\underline{c}}^{c^{B1}} m_1 c \mathrm{d}G(c) + \int_{\underline{c}}^{c^{B2}} m_2 c \mathrm{d}G(c) = B$. 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.

prices there are always equal to marginal costs. In Grassi and Ma (2008), we let consumer wealth be a continuous random variable, and derived the optimal rationing policy when the private market is competitive. The results of that analysis applies here, where the distribution of consumers wealth is discrete. In the optimal scheme, all consumers are rationed when their costs are above a threshold. When costs are below the threshold, some consumers with lower wealth will be assigned the good for free, while others with higher wealth will again be rationed. Given a competitive market, the public supplier is no longer concerned with implementing price reduction. When consumers have low costs, they derive high inframarginal gain from the competitive private market. Moreover, richer consumers enjoy higher inframarginal surplus. The public supplier may save resources on low-cost consumers, let the more wealthy buy from the private market, and use the released resources to supply those with higher costs. It is of interest to note that given a competitive private market, the optimal rationing rule uses both cost and wealth information, whereas when the private market is a monopoly, optimal rationing uses only cost information.[3]

5 Concluding remarks

We have presented a model to study the effect of rationing on prices in the private market. Public policies should take into account market responses. We show 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; wealth information is not necessary. Our model sheds light on crowd out, and the design of public programs when private market responses are important.

We assume two wealth classes to make the model tractable. Extending the model and deriving the equilibrium rationing scheme for many, or a continuum of wealth classes involve complex computation, because many possible price reductions must be considered. We believe, however,

[3] This result is consistent with Hoel (2007). In his model, there are many goods, which vary in terms of benefits and costs. Hoel assumes a competitive private market.

that our insight will be robust. We have used a separable utility assumption. In Grassi and Ma (2008), we list some factors that may influence the results when utility functions are not separable in money and benefits from the good. A secondary effect from the consumption of the good on the marginal utility of income will have to be considered. Again, for consumption of goods that will yield low income effects, our results will extend to the more general utility functions.

We have assumed a fixed budget. Extending the model to consider an optimal budget is fairly straightforward. We have obtained the optimal policies in the two information regimes, so that the optimized aggregate consumer utility is available. Once the cost of public funds is specified, the usual optimization steps can be taken to characterize the optimal budget.

The analysis here is limited to free public supply. A fixed user fee can be incorporated into the model without altering any results. Due to risk aversion, publicly provided health insurance usually does not impose significant copayments. The same is true in the education market, because low-income households may be cash-constrained. Nevertheless, a general analysis of optimal monetary subsidy may be fruitful.

Appendix

Proof of Proposition 1: Because all terms in square brackets in the objective function (7) are constant, we alternatively can write the objective function as $m_2\theta_2 G(c_1)$.

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 $m_2\theta_2 G(c_1) = 0$. 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 $m_2 k G(c_m) > 0$. This implies that at a solution $c_1 > \underline{c}$ and $\theta_2 > 0$. Because $c_1 > \underline{c}$, it follows from (4) that θ_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

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

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

$$m_2\theta_2 = K \frac{\tau_1 - c_1}{\tau_2 - c_1}. \tag{22}$$

We next substitute (22) into the objective function $m_2\theta_2 G(c_1)$. The constrained maximization problem (with the boundary conditions $\theta_i \leq 1$ omitted) is the same as the unconstrained maximization problem:

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

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

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

The left-hand side of (23) is increasing in c_1. For c_1 between $\underline{c}$ and τ_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-hand of (23) is zero. Therefore, there exists a unique c_1^* strictly between $\underline{c}$ and τ_1 that satisfies (23).

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

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

If β 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 β 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 τ_1 or τ_2 depending on whether (10) is satisfied. If we have defined p by the method just 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 (ϕ_1, ϕ_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 λ denote the multiplier for the budget constraint (15), and $\mu(c)$ the multiplier for (16) at c. The Lagrangean is

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

where we have omitted the boundary conditions on ϕ_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

$$\frac{\partial L}{\partial \phi_1} = -m_1 + \lambda m_1 c \tag{25}$$

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

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

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

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

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

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 τ_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 τ_1, $c_m < c^B < \tau_1$. Recall that c_m is the cost level at which the firm will set the low price τ_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

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

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 $p(c) = 0$, 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 by the same argument 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 $\phi_i(c) = 1$, 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 MasColell , "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.

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