PUBLIC SECTOR RATIONING AND PRIVATE SECTOR SELECTION

SIMONA GRASSI
University of Lausanne

CHING-TO ALBERT MA
Boston University

Abstract

We study the interaction between nonprice public rationing and prices in the private market. Under a limited budget, the public supplier uses a rationing policy. A private firm may supply the good to those consumers who are rationed by the public system. Consumers have different amounts of wealth, and costs of providing the good to them vary. We consider two regimes. First, the public supplier observes consumers' wealth information; second, the public supplier observes both wealth and cost information. The public supplier chooses a rationing policy, and, simultaneously, the private firm, observing only cost but not wealth information, chooses a pricing policy. In the first regime, there is a continuum of equilibria. The Pareto dominant equilibrium is a means-test equilibrium: poor consumers are supplied while rich consumers are rationed. Prices in the private market increase with the budget. In the second regime, there is a unique equilibrium. This exhibits a cost-effectiveness rationing rule; consumers are supplied if and only if their cost–benefit ratios are low. Prices in the private market do not change with the budget. Equilibrium consumer utility is

> higher in the cost-effectiveness equilibrium than the means-test equilibrium.

1. Introduction

Many governments and public organizations provide or subsidize goods and services such as health care and education. Free or subsidized public provision often coexists with a private market. In this paper, we study the interaction between rationing policies in the public sector and profit-maximizing prices in the private sector. We derive equilibria of games between the public and private sectors, and compare equilibrium prices and aggregate consumer utilities.

All public programs operate under limited budgets. Unable to cover all intended consumers, a public supplier must use a rationing rule. A variety of rationing and subsidy practices exist. Current Medicaid policies in the United States provide health insurance to indigent individuals. In the pending U.S. health care reform, families up to 400% of the Federal Poverty Level receive either free health insurance or substantial subsidies. These rationing policies are a means-test mechanism, which allocates public supply to poor individuals. In Canada and many European countries, public health systems ration care according to illnesses, patients' medical conditions, and treatment costs. This form of rationing is based on a cost-benefit or cost-effectiveness mechanism, which allocates public supply to individuals for whom it is worthwhile, in some qualified sense.

Rationed individuals nevertheless may consider purchasing from the private market. Here, an important issue is selection of profitable consumers by a private firm. A focus of our paper is on how a private firm's profit-maximizing strategy will react to the public rationing mechanism. For example, when an individual does not qualify for Medicaid, his income should not be very low; neither should be his willingness to pay for health insurance. On the other hand, if a patient does not qualify for a certain treatment according to the public sector cost-effectiveness measure, a private firm may not infer about his willingness to pay.

Our model consists of a set of consumers, a public supplier, and a private firm. Each consumer would like at most one unit of an indivisible good (a medical treatment, a course of education, etc.). Consumers have different wealth or income levels, and the costs of providing the good to them also differ. We use wealth heterogeneity to model differences in consumers' valuations of the good. Rich consumers are more willing to pay for the good than poor consumers. Cost heterogeneity arises because a consumer's characteristics may determine how much it costs to supply the good to him. For example, the cost of a medical treatment depends on illness severity, and the cost of helping a student to achieve an academic standard depends on the student's ability and aptitude. Variations in these characteristics affect provision costs.

To focus on rationing and pricing issues, we abstract from financing. The public supplier has a limited budget for providing the good at zero cost to consumers, and aims to maximize total consumer utility.[1] We consider two information regimes. In the first, rationing is based on consumers' wealth information. In the second, rationing is based on consumers' wealth and cost information. The first regime corresponds to the means-test scenario above. The second regime is somewhat more general than the cost-effectiveness scenario, but we will exhibit a unique equilibrium in which the public supplier ignores wealth information, so cost-effectiveness does obtain.

We study the equilibria of the following games. In stage 1, nature draws randomly and independently the wealth level and cost for each consumer. Consumers' cost information is learned by the firm. In the first regime, the public supplier learns consumers' wealth information, but not the costs. In the second regime, the public supplier learns both consumers' wealth and cost information. In stage 2, the public supplier designs a rationing scheme based on the available information in each regime. Simultaneously, the private firm sets a price schedule depending on costs. In stage 3, the public supplier's rationing scheme is implemented, and rationed consumers decide whether to purchase from the firm at prices set in the second stage. We model the private sector by a monopoly firm, and the analysis extends to Cournot competition.

Our model addresses the following issues. First, the private sector may react to public supply by selecting or cream-skimming consumers. How does the public sector react to cream-skimming? Second, different consumers enjoy different surpluses from public and private supplies. How does this affect public supply and prices in the private sector? Third, public rationing policies may be based on different information. How do equilibria change as the information structure changes?

In the first regime, when public rationing is based on consumers' wealth, there is a continuum of equilibria. In the Pareto dominant equilibrium, the public supplier uses the budget on poor consumers and rations all rich consumers. The means-test criterion emerges in equilibrium. Prices set by the private firm rises as the available budget increases. As poor consumers are supplied, the private firm realizes that it sells only to consumers with higher willingness to pay, so raises prices accordingly.

In the second regime, when public rationing is based on consumers' wealth and cost, there is a unique equilibrium. The public supplier spends the entire budget on consumers whose cost per unit of benefit is low,

_____

[1] Nonprice rationing is ubiquitous. Many governments set negligible prices for public education and health care. Presumably, this may be due to fairness or political considerations. There are also economic reasons for favoring rationing. For the health market, insuring consumers' financial risks due to illness is a fundamental goal. Under social insurance, consumers should not be exposed to too much financial risk upon becoming sick. For the education market, a government may encourage the investment of human capital, which may enhance economic growth and create externalities. Again, reducing costs of education may be a sensible policy.

and ignores consumers' wealth information. The cost-effectiveness criterion emerges in equilibrium. The private and public sectors will appear to be separated, with the public sector serving consumers with low cost-benefit ratios, while the private sector serving those with high cost–benefit ratios. Equilibrium private market prices for available consumers remain the same as if the public sector were inactive, and equilibrium prices are independent of the budget. At any given budget level, equilibrium aggregate consumer utility is higher when rationing is based on the cost-effectiveness criterion than the means-test criterion.

What explains these results? Richer consumers have higher willingness to pay than poor consumers. In the private market, consumers face prices that are dependent on their costs, not wealth. Hence, richer consumers obtain higher incremental surplus from trades in the private sector. This trade-surplus effect is common across both information regimes. When rationing is based on wealth, releasing richer consumers to the private market allows them to realize more incremental trade surplus there than poorer consumers. This motivates the public supplier to ration the rich. Simultaneously, when public supply is for poor consumers, the firm knows that only richer consumers are in the market. These consumers have higher willingness to pay, so the firm raises prices.

In the second information regime, cost information is available. Here, absent the private sector, a cost-effectiveness principle applies: public supply is for consumers with a low cost–benefit ratio. The cost-effectiveness principle continues to apply when there is a private market. For a given cost level, a consumer obtains more surplus from free public supply than the private market. Rationing low-cost and rich consumers cannot occur in equilibrium. If such consumers were rationed, the private firm would understand that low-cost consumers must be rich, and would raise prices accordingly. The trade-surplus effect cannot be implemented.

Our goal is to study the effect of a price-reactive private sector, so we have rejected a perfectly competitive market where prices would always be marginal costs (but for completeness, we have included related results). We use the monopoly setup, but all results extend to Cournot competition. We have let the monopolist observe costs, but not wealth. If the firm would also observe wealth, it would be able to extract all surplus, and the public rationing policy would not affect consumers' trade surplus in the private market. In any case, firms seldom possess information on wealth. Given that the good is indivisible and a consumer buys at most one unit, nonlinear prices cannot be implemented. It would be a less interesting model if the firm did not observe consumers' costs; selection and cream-skimming issues would be assumed away.[2]

Most other papers assume that the government is a first mover. Generally, an ability to commit to a rationing rule is valuable. However, as we show

---

[2] Also, when the price is based on the expected cost, the firm may renege once it learns that the cost turns out to be higher than the price it has charged.

in a companion paper, Grassi and Ma (2011), the public supplier's Stackelberg rationing rule is time-inconsistent. By committing to ration some poor consumers, the public supplier can implement lower prices because the firm will want to sell to poor consumers when costs are low. However, given lower prices, the public supplier would exploit the trade-surplus effect by reneging, supplying poor consumers and rationing rich consumers. We have adopted a simultaneous-move game, and the public supplier and the private firm have symmetric commitment power. This is a long-term perspective between the players because both players choose mutual best responses. Finally, our focus is on nonprice rationing. Free public provision is the main mechanism in public health and education systems, so we have not included taxes and subsidies in the public supplier's strategies.

We study interactions between public rationing and a noncompetitive private market, while models in the literature usually assume a competitive private market, or exogenous pricing rules. Barros and Olivella (2005) focus on public physicians referring patients to their own private practices. The public sector uses waiting-time rationing, and physicians refer patients when the patients' costs are low. Iversen (1997) considers the effect of a private sector on waiting time in the public sector. Hoel and Sæther (2003) consider supplementary private health care when public health care is subject to waiting time rationing. In the above papers, the price in the private market is fixed. Hoel (2007) derives the optimal cost-effectiveness rule when patients have access to a competitive private market, where prices do not respond to the public sector's allocation rule. By contrast, in our model, prices in the private sector respond to public policies.

Rationing is similar to transfers in kind. Blackorby and Donaldson (1988) show how transfers in kind may solve asymmetric information problems. The literature has also studied redistribution effects. In Besley and Coate (1991), the government uses a poll tax to provide for free a good at a low quality. Rich consumers optimally choose the good at a high quality from a competitive private market, while poor consumers do not. The government in effect taxes rich consumers to subsidize poor ones. Segregation between rich and poor consumers is an equilibrium of our game when rationing is based on wealth, but this stems from a trade-surplus effect in the private market. In Besley and Coate (1991) the budget is endogenous while we abstract from financing issues, and assume an exogenous budget.

Our formal model is like a common agency model. The public supplier and the private firm are two principals whose actions will affect the consumer, who is the agent; see Bernheim and Whinston (1986). In line with the common agency model, we use a symmetric setup, so that both sectors react against each other's strategy. We are unaware of a paper that models how public sector rationing and private sector pricing strategies mutually react. We also depart from the mechanism design literature on the provision of public goods (see Norman 2004, and Hellwig 2003 on excludable public goods), where incentive-compatible, individually rational and budget-balanced schemes are derived. In Norman (2004) the public supplier may set

individualized user fees to provide access to the excludable public good. Our interest is on nonprice rationing, so we have not allowed the public supplier to use taxes and subsidies.

Section 2 and its subsections lay out the model. In Section 3, we first derive a focal equilibrium in the wealth rationing regime. Then we characterize a continuum of other equilibria. Section 4 contains the analysis of rationing based on wealth and cost, and compares equilibrium consumer utilities across the two rationing regimes. The last section draws some conclusions. Appendix A contains proofs of all lemmas, propositions, and the last corollary; proofs of other corollaries are omitted. Appendix B discusses alternative assumptions on consumer preferences, information, and cost–benefit structures.

2. The Model

We begin with the description of consumers. Next, we introduce a public sector, and derive benchmark optimal rationing policies. Then we introduce a private firm. We complete the model by describing the extensive-form games between consumers, the public supplier, and the private firm.

2.1. Consumers and Their Willingness to Pay

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

Let $F : [\underline{w}, \overline{w}] \to [0, 1]$ be the distribution function of w. We assume that F is differentiable, and the corresponding density f strictly positive. Similarly, let $G : [\underline{c}, \overline{c}] \to [0, 1]$ be the distribution function of c. We also assume that G is differentiable, and the corresponding density g strictly positive. Let $\gamma \equiv \int c \, \mathrm{d}G$ denote the expected cost. The domains of both distributions are strictly positive and bounded. The variables w and c are assumed to be independently distributed. In Appendix B, we will discuss the independence assumption.

For a general specification of preferences, we can let a consumer's utility be $U(w, 0)$ when he does not consume the good, and $U(w - p, 1)$ when he consumes the good at a price $p \geq 0$. The utility function U is strictly increasing, and strictly concave in w, and $U(w, 1) > U(w, 0)$. It saves on notation and simplifies the analysis if we let the utility function U be separable in the two arguments. The separability assumption says that a unit of the good

generates the same utility increment, independent of the consumer's wealth. If the consumer with wealth w pays a price p to consume the good, his utility is $U(w - p) + 1$; his utility is $U(w)$ if he does not consume the good. We discuss the nonseparable utility function in Appendix B.

If a consumer with wealth w is indifferent between paying τ for the good and the status quo, we have

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

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

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

A consumer's willingness to pay for the good is strictly increasing in wealth due to the strict concavity of U. We will assume that the lowest willingness to pay $\tau(\underline{w})$ is larger than the lowest cost $\underline{c}$. This assumption ensures that there is some scope for any consumer to benefit from a trade in the private market.

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

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

2.2. The Public Sector and Rationing

The public sector has a budget B which is insufficient to supply the good to all consumers for free, so $0 < B < \gamma$. We consider two information regimes. First, only consumers' wealth information is available to the public supplier, and second, consumers' wealth and cost information is available. In each case, nonprice rationing will be used to allocate the budget for providing the good to consumers.

In the first regime, the public supplier's rationing rule is a function $\theta : [\underline{w}, \overline{w}] \to [0, 1]$. For $w \in [\underline{w}, \overline{w}]$, the public supplier provides consumers

with wealth below w a total of $\int_{\underline{w}}^{w} (1 - \theta(x)) f(x)\, dx$ units of the good. The rationing rule θ splits the density f so that at w, $[1 - \theta(w)] f(w)$ of consumers are supplied at zero price, and $\theta(w) f(w)$ of consumers are rationed. Because wealth and cost are independently distributed, the cost c among rationed consumers remains distributed according to G.

In the second regime the rationing rule is a function $\phi : [\underline{w}, \overline{w}] \times [\underline{c}, \overline{c}] \to [0, 1]$. It has the same interpretation as in the first regime. For consumer (w, c), the density $\phi(w, c) f(w) g(c)$ is available to the private firm.[3] In each regime, the public supplier's objective is the sum of consumer utilities.

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

2.3. Benchmark Optimal Rationing Policies with an Inactive Private Market

For now suppose that the public sector is the sole provider. Consider the first information regime where rationing is based on wealth. For a rationing rule θ, total consumer benefit from the public supply is $\int_{\underline{w}}^{\overline{w}} (1 - \theta(w)) f(w)\, dw$ as each unit of consumption increases a consumer's utility by one unit. The consumer welfare index, which the public supplier maximizes, is

$$V(\theta) \equiv \int_{\underline{w}}^{\overline{w}} U(w)\, dF + \int_{\underline{w}}^{\overline{w}} [1 - \theta(w)]\, f(w)\, dw. \tag{3}$$

The rationing rule must satisfy the budget constraint

$$\gamma \int_{\underline{w}}^{\overline{w}} [1 - \theta(w)] f(w)\, dw \leq B, \tag{4}$$

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

---

[3] We restrict rationing rules to those that leave the functions $\theta(w) f(w)$ and $\phi(w, c) f(w)$ integrable, so that $\int_{\underline{w}}^{w} \theta(x) f(x)\, dx$ and $\int_{\underline{w}}^{w} \phi(x, c) f(x)\, dx$ at each c are well defined for $w \in [\underline{w}, \overline{w}]$. We can restrict the public provider to supply to either all or none of the consumers within a wealth class or a wealth-cost class. Rationing schemes are then functions that map $[\underline{w}, \overline{w}]$ to $\{0, 1\}$ and $[\underline{w}, \overline{w}] \times [\underline{c}, \overline{c}]$ to $\{0, 1\}$. The general rationing functions can now be interpreted as mixed strategies. For ease of exposition, we do not use the mixed strategy interpretation.

The determination of a rationing rule that maximizes (3) subject to (4) is rather trivial. Any rationing rule that exhausts the budget is optimal. The public supplier allocates the good to consumers without collecting any payment. Due to the separable utility function, the utility increment is independent of w. Any rationing scheme that exhausts the budget results in the same level of the welfare index, and is optimal.

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

$$V(\phi) \equiv \int_{\underline{c}}^{\overline{c}} \int_{\underline{w}}^{\overline{w}} \{U(w) + [1 - \phi(w, c)]\} \, f(w) g(c) \, \mathrm{d}w \, \mathrm{d}c. \qquad (5)$$

The rationing rule must satisfy the budget constraint

$$\int_{\underline{c}}^{\overline{c}} \int_{\underline{w}}^{\overline{w}} [1 - \phi(w, c)] c \, f(w) g(c) \, \mathrm{d}w \, \mathrm{d}c \leq B. \qquad (6)$$

By pointwise optimization with respect to ϕ, the optimal rationing rule is given by[4]

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

where $\int_{\underline{c}}^{c^s} c \, \mathrm{d}G(c) = B$.

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

2.4. The Private Sector and Profits

There is a monopoly firm in the private sector; Cournot and perfect competition in the private sector will be discussed. For consumer (w, c), the private firm observes the cost of providing a unit of the good to the consumer c, but not his wealth w.

By setting a price p, the monopolist sells to those consumers with willingness to pay higher than p. Obviously, the monopolist will not set a price outside the range of willingness to pay τ. Setting a price p is equivalent to selecting the wealth level of the *marginal* consumer w, where $p = \tau(w)$. By the strictly monotonicity of τ, consumers with $w' > w$ have $\tau(w') > \tau(w)$, hence are willing to pay $p = \tau(w)$ to purchase the good. The function τ is like a demand function; we simply restate the common principle that a monopolist may choose equivalently between a price and a quantity while respecting the demand function. It is more convenient to let the firm choose the quantity

---

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

or the marginal consumer w by setting the price $\tau(w)$. A quantity function is denoted by $\widehat{w} : [\underline{c}, \overline{c}] \rightarrow [\underline{w}, \overline{w}]$.

We present the profit functions under the two rationing rules. First, suppose that rationing is based on wealth. At cost c the density of consumers available to the firm is $\theta(w) f(w)$. At a price $\tau(w)$, consumers with wealth higher than w will buy, and the profit is

$$\pi(w; c, \theta) = \int_{w}^{\overline{w}} \theta(x) f(x) \, \mathrm{d}x \, [\tau(w) - c]. \tag{7}$$

Here, the integral is the total quantity purchased, and $\tau(w) - c$ is the price-cost margin. Second, suppose that rationing is based on wealth and cost. At cost c the density of consumers available to the firm is $\phi(w, c) f(w)$. At a price $\tau(w)$, consumers with wealth higher than w will buy, and the profit is

$$\pi(w; c, \phi) = \int_{w}^{\overline{w}} \phi(x, c) f(x) \, \mathrm{d}x \, [\tau(w) - c]. \tag{8}$$

2.5. Interaction between the Public and Private Sectors

We study the following games and look for their subgame-perfect equilibria:

Stage 1: Nature draws (w, c) according to the distributions F and G, respectively, for each consumer. The private firm observes c. The public supplier observes either w, or both w and c.

Stage 2: In each information regime, the public supplier chooses a rationing rule, θ or ϕ, and the private firm chooses a quantity function $\widehat{w}$.

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

3. Equilibrium Rationing and Prices When Rationing is Based on Wealth

We begin with the private firm's profit-maximizing prices and consumers' utilities. Then we present an equilibrium in which the public supplier uses the entire budget on consumers with lower wealth levels. Next, we present a continuum of equilibria, and show that the one we have presented is Pareto dominant. Finally, we discuss some comparative statics on the budget, as well as Cournot and perfect competition in the private market.

3.1. Profit-Maximizing Prices and Consumer Utilities

To characterize an equilibrium, we need to refer to a profit-maximizing quantity function when the firm has access to all consumers. Let this function
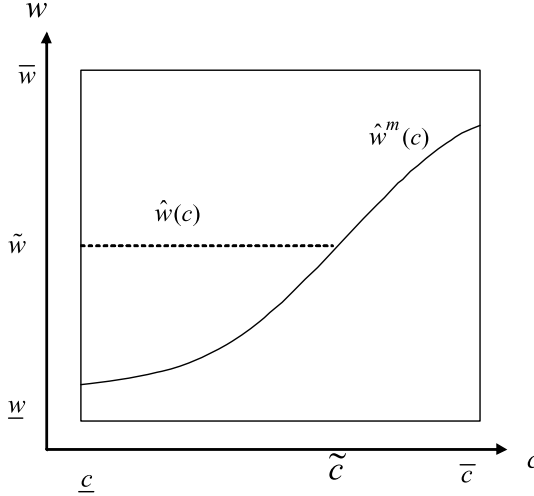
Figure 1: Quantity functions $\widehat{w}^m(c)$ and $\widehat{w}(c)$.

be $\widehat{w}^m(c)$. Suppose that the public supplier rations all consumers, so in (7) we set $\theta(w) = 1$ for all w. The function $\widehat{w}^m : [\underline{c}, \overline{c}] \to [\underline{w}, \overline{w}]$ is given by

$$\widehat{w}^m(c) \equiv \operatorname*{argmax}_{w} \int_{w}^{\overline{w}} f(x)\, \mathrm{d}x\, [\tau(w) - c]. \tag{9}$$

We assume that the profit function in (9) is concave, and that $\widehat{w}^m(c)$ is single-valued. By the Maximum Theorem $\widehat{w}^m(c)$ is continuous. We further assume that as c varies over $[\underline{c}, \overline{c}]$, the marginal consumers vary over a proper subset of $[\underline{w}, \overline{w}]$, so $\underline{w} < \widehat{w}^m(\underline{c}) < \widehat{w}^m(\overline{c}) < \overline{w}$. This requires that variation in wealth is sufficiently large relative to variation in costs. The optimal quantity $\widehat{w}^m(c)$ is given by the usual marginal-revenue-equal-marginal-cost condition. The quantity function $\widehat{w}^m(c)$ is strictly increasing. As marginal cost increases, the optimal quantity is adjusted to achieve a higher level of marginal revenue. This means setting a higher price and selling to less consumers.[5] By the concavity of the profit function for any $w > \widehat{w}^m(c)$, the derivative of profit with respect to w is negative:

$$\frac{\mathrm{d}}{\mathrm{d}w} \int_{w}^{\overline{w}} f(x)\, dx\, [\tau(w) - c] < 0 \quad \text{for} \quad w > \widehat{w}^m(c), \quad \text{any } c. \tag{10}$$

The solid line in Figure 1 illustrates such a quantity function. We have assumed that the budget is insufficient to cover all consumers at zero cost $(B < \gamma)$. For a very small budget, the interaction between the public and

---

[5] The sign of the derivative of $\widehat{w}^m(c)$ is the same as the sign of the cross partial of $\int_{w}^{\overline{w}} f(x)\, \mathrm{d}x\, [\tau(w) - c]$, which is positive.

private sector may be irrelevant. Suppose that the budget can only cover some consumers with $w < \widehat{w}^m(\underline{c})$. The private firm will never sell to these consumers even if they are available because their willingness to pay is too low. Any public supply to them will not lead to a price response from the private firm. To rule out equilibria without any interaction between sectors, we assume that the budget is sufficient to cover some consumers who may purchase from the firm. This assumption is $B > F(\widehat{w}^m(\underline{c}))\gamma$. At the lowest cost $\underline{c}$, the firm sells to those with wealth above $\widehat{w}^m(\underline{c})$, so this assumption says that the budget is sufficient to eliminate some consumers who otherwise may buy from the firm.

Given a rationing rule θ, the private firm's profit from selling to consumers with wealth higher than w is in (7). Let $\widehat{w}(c)$ be the optimal quantities, and $\widehat{\pi}(c)$ the maximum profit:

$$\widehat{w}(c) \equiv \arg\max_{w} \pi(w; c, \theta), \tag{11}$$

$$\widehat{\pi}(c) \equiv \pi(w'; c, \theta), \quad w' \in \widehat{w}(c). \tag{12}$$

For some rationing rules, there may be multiple quantities that maximize profit, so $\widehat{w}(c)$ is a correspondence. According to the Maximum Theorem, the correspondence $\widehat{w}(c)$ is upper semicontinuous. The profit-maximizing price $\tau(\widehat{w}(c))$ may not be strictly increasing in cost c, although the maximum profit is strictly decreasing, as the next lemma shows.

LEMMA 1: *The maximum profit is strictly decreasing in c. Any selection from the profit-maximizing quantities $\widehat{w}(c) \equiv \arg\max_w \pi(w; c, \theta)$ is increasing in c. That is, if $c_1 < c_2$, then $w_1 \leq w_2$, where $w_1 \in \widehat{w}(c_1)$ and $w_2 \in \widehat{w}(c_2)$.*

A best response is a selection from the profit-maximizing quantity correspondence, and need not be continuous. Nevertheless, because it must be (weakly) increasing, any point of discontinuity of $\widehat{w}(c)$ must be an upward jump (see also Figure 3). Because there is no risk of confusion, we also denote such a (weakly increasing) selection by the notation $\widehat{w} : [\underline{c}, \overline{c}] \to [\underline{w}, \overline{w}]$.

When will an equilibrium quantity function fail to be strictly increasing? Suppose that the rationing rule specifies that $\theta(w) = 0$ for $w < \widetilde{w}$, and $\theta(w) = 1$ for $w > \widetilde{w}$. This rationing scheme supplies consumers (at zero price) if and only if their wealth is below a threshold $\widetilde{w}$; consumers with $w < \widetilde{w}$ are not in the market. Let the cost threshold $\widetilde{c}$ be defined by $\widehat{w}^m(\widetilde{c}) = \widetilde{w}$; see Figure 1. For $c > \widetilde{c}$, the profit-maximizing quantity remains at $\widehat{w}^m(c)$. At $c > \widetilde{c}$, the firm would not sell to those consumers with wealth below $\widetilde{w}$ anyway. At $c < \widetilde{c}$, the profit-maximizing quantity function is constant at $\widetilde{w}$. Lemma 1 says that the profit-maximizing quantity cannot rise as cost falls below $\widetilde{c}$. The firm has no available consumers with wealth below $\widetilde{w}$, so the
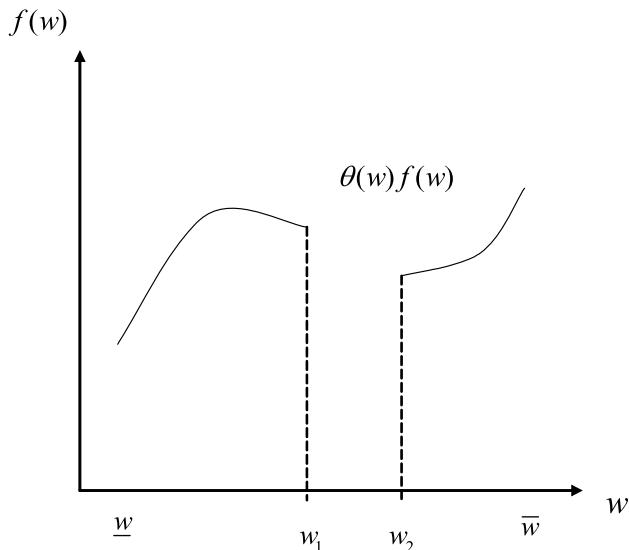
Figure 2: Consumer density under rationing scheme θ.

optimal quantity stays at $\widetilde{w}$.[6] Under the rationing scheme, in Figure 1 the optimal quantity function becomes the horizontal, dotted line when cost falls below $\widetilde{c}$.

When will an equilibrium quantity function fail to be continuous? Suppose that $\theta(w) = 0$ for $w \in [w_1, w_2]$ where $\underline{w} < w_1 < w_2 < \overline{w}$, and $\theta(w) = 1$ otherwise. The public sector supplies only to consumers with medium wealth. Figure 2 illustrates the density of consumers available to the private firm. The profit-maximizing quantity function $\widehat{w}(c)$ is in Figure 3. For $c < c_1$ or $c > c_2$, the profit-maximizing quantity is unique. For $c \in (c_1, c_2)$, the price remains constant because all consumers with wealth in $[w_1, w_2]$ are supplied by the public sector. At cost c_1, the firm makes equal amounts of profit whether it charges a price $\tau(w_2)$ selling to consumers with $w > w_2$, or $\tau(w_0)$ selling to consumers with w between w_0 and w_1 and above w_2. Finally, some quantities may never be chosen; in Figure 3, the firm never sets $\widehat{w}(c)$ to any $w \in [w_0, w_2)$.

We need to write down aggregate consumer utility given rationing functions and quantity functions that are weakly increasing and possibly

---

[6] Any $w < \widetilde{w}$ yields a profit

$$\left[\int_w^{\widetilde{w}} \theta(x) f(x)\, \mathrm{d}x + \int_{\widetilde{w}}^{\overline{w}} \theta(x) f(x)\, \mathrm{d}x \right] [\tau(w) - c]$$

$$= \left[\int_{\widetilde{w}}^{\overline{w}} f(x)\, \mathrm{d}x \right] [\tau(w) - c] < \left[\int_{\widetilde{w}}^{\overline{w}} f(x)\, \mathrm{d}x \right] [\tau(\widetilde{w}) - c]$$

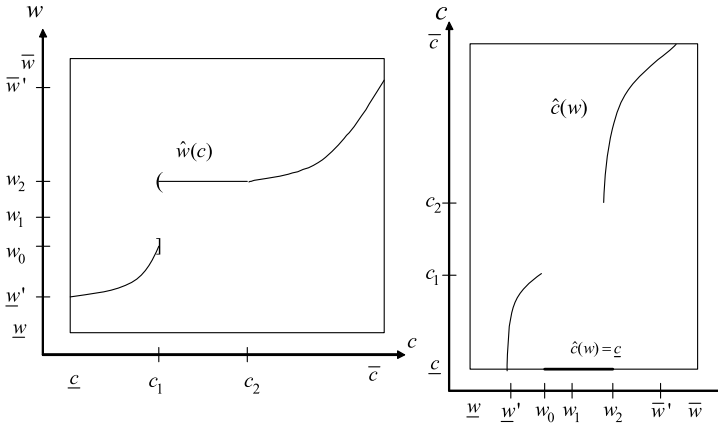so setting $w = \widetilde{w}$ is optimal.

Figure 3: The quantity function $\widehat{w}(c)$ and its "inverse" $\widehat{c}(w)$.

exhibiting upward jumps. Given a quantity function, $\widehat{w}(c)$, rationed con-
sumer (w, c) buys from the private firm if and only if $w \geq \widehat{w}(c)$. In Figure 3,
this is the set above the graph of $\widehat{w}(c)$. It is more convenient to view the set
of purchasing consumers as one indexed by a function $\widehat{c} : [\underline{w}, \overline{w}] \to [\underline{c}, \overline{c}]$
that is like an "inverse" of $\widehat{w}$. Define $\widehat{c}(w) = \sup\{c : w \geq \widehat{w}(c)\}$; if there is no
$c \in [\underline{c}, \overline{c}]$ such that $w \geq \widehat{w}(c)$, set $\widehat{c}(w) = \underline{c}$. Such a function $\widehat{c}$ is illustrated in
Figure 3 .

While the function $\widehat{w}$ gives the wealth of the marginal consumer in terms
of his cost, the function $\widehat{c}$ gives the threshold cost level below which a con-
sumer with wealth w will buy at price $\tau(\widehat{w}(c))$. Whenever $\widehat{w}$ is strictly increas-
ing and continuous, the function $\widehat{c}$ is its inverse. When $\widehat{w}$ is constant on an
interval, then $\widehat{c}$ exhibits discontinuities at the two ends of the interval. Finally,
$\widehat{c}(w)$ becomes $\underline{c}$ when the firm does not sell to consumer (w, c). Clearly $\widehat{c}(w)$
is increasing whenever its value is not $\underline{c}$. The set of consumers who purchase
are those with (w, c) below the graph of $\widehat{c}(w)$, and this differs from those
above the graph of $\widehat{w}(c)$ at most for a set of measure zero. Functions $\widehat{w}$ and
$\widehat{c}$ are two equivalent ways of keeping track of consumer types who purchase
from the private firm.

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

$$\int_{\underline{w}}^{\overline{w}} [1 - \theta(w)] f(w) [U(w) + 1] \, dw$$

$$+ \int_{\underline{w}}^{\overline{w}} \theta(w) f(w) \left[\begin{array}{l} \displaystyle\int_{\underline{c}}^{\widehat{c}(w)} \{U(w - \tau(\widehat{w}(c))) + 1\} g(c) \, dc \\[2ex] + \displaystyle\int_{\widehat{c}(w)}^{\overline{c}} U(w) g(c) \, dc \end{array} \right] \, dw. \qquad (13)$$

In this expression, the first term is the utility of consumers supplied by the public sector. The second term is the utility of rationed consumers. The top integral inside the big square brackets is the utility of consumer (w, c) buying from the firm at price $\tau(\widehat{w}(c))$, while the bottom integral is the utility of consumers who do not buy. The welfare index can be simplified to

$$
\begin{aligned}
V(\theta) = & \int_{\underline{w}}^{\overline{w}} \left[U(w) + (1 - \theta(w)) \right] f(w) \, \mathrm{d}w \\
& + \int_{\underline{w}}^{\overline{w}} \theta(w) f(w) \left[\int_{\underline{c}}^{\widehat{c}(w)} \{ U(w - \tau(\widehat{w}(c))) + 1 - U(w) \} g(c) \, \mathrm{d}c \right] \mathrm{d}w,
\end{aligned}
$$
$$(14)$$

where the first term is the base utility $U(w)$ plus the increase of utility from public supply, and the second term is the consumer incremental surplus from the private market.

3.2. An Equilibrium: Rationing Rich Consumers

An equilibrium is a pair of rationing and quantity functions $(\theta, \widehat{w})$ that are mutual best responses. That is, θ maximizes aggregate consumer utility subject to the budget constraint given quantity function $\widehat{w}$, and $\widehat{w}$ maximizes profit for every c given θ. We first present an equilibrium in which consumers are rationed if and only if they are rich, and in which the firm chooses the quantity function $\widehat{w}(c)$ in Figure 1. This is a means-test equilibrium: consumers are supplied if and only if their wealth is low.

PROPOSITION 1:   *The following is an equilibrium. The public supplier rations all consumers with wealth above a threshold w^E and supplies all consumers with wealth below w^E: $\theta(w) = 1$, $w > w^E$ and $\theta(w) = 0$, $w < w^E$. The threshold w^E exhausts the budget and is given by $F(w^E)\gamma = B$. The private firm sets the price to the monopoly price when cost is above a threshold c^E, defined by $\widehat{w}^m(c^E) = w^E$, and to a fixed price $\tau(w^E)$ when cost is below c^E.*

We explain the intuition for Proposition 1. First, when discussing the property of $\widehat{w}(c)$ in Figure 1, we already show that it is the profit-maximizing quantity function when the rationing policy is the one in Proposition 1. Hence, the firm's quantity function in Proposition 1 is a best response.

We now explain why the rationing policy is a best response. To compute aggregate consumer utility, we apply the quantity function in the Proposition to (14). First, for $c > c^E$ we let the inverse of $\widehat{w}^m$ be $\widehat{c}^m$; this takes the role of $\widehat{c}$ in (14): for $w > w^E$, consumer (w, c) buys at $\tau(\widehat{w}^m(c))$ if $c < \widehat{c}^m(w)$. Second, for $c < c^E$, $\widehat{c}^m$ takes the value $\underline{c}$: rationed consumers with wealth below w^E never buy because the firm never sets a price below $\tau(w^E)$.

Under rationing policy θ, aggregate consumer utility is

$$\int_{\underline{w}}^{\overline{w}} \left[U(w) + (1 - \theta(w)) \right] f(w) \, \mathrm{d}w$$

$$+ \int_{w^E}^{\overline{w}} \theta(w) \left\{ \begin{array}{l} \int_{\underline{c}}^{c^E} \left[U(w - \tau(w^E)) + 1 - U(w) \right] g(c) \, \mathrm{d}c \\[2ex] + \int_{c^E}^{\widehat{c}^m(w)} \left[U(w - \tau(\widehat{w}^m(c))) + 1 - U(w) \right] g(c) \, \mathrm{d}c \end{array} \right\} f(w) \, \mathrm{d}w. \tag{15}$$

The welfare index is described as follows. The integral on the first line of (15) is the sum of the base utility $U(w)$ plus the utility increase from the public supply. The second line is the private market incremental surplus of rationed consumers. Only rationed consumers with wealth above w^E will buy from the private market. For these consumers, if their costs are below c^E, they purchase at price $\tau(w^E)$, and obtain the incremental surplus in the integral with limits between $\underline{c}$ and c^E; if their costs are above c^E, they purchase at price $\tau(\widehat{w}^m(c))$ if their costs are below $\widehat{c}^m(w)$, and obtain the incremental surplus in the integral with limits between c^E and $\widehat{c}^m(w)$.

A rationing policy θ is a best response if it maximizes (15) subject to the budget constraint (4). We consider the trade-off in rationing a consumer with wealth w. The benefit of rationing a consumer is the saving of expected cost γ, a constant. The cost of rationing a consumer depends on the consumer's wealth level. If w is below w^E, this consumer does not buy from the private market, so the cost is one unit of utility due to nonconsumption. If w is above w^E, this consumer may gain some incremental surplus from the private market (the second line in (15)), so the cost of rationing him is less than one unit of utility. Rationing a rich consumer is less costly precisely because the rich consumer has the opportunity to buy from the private market. Therefore, it is optimal to ration richer consumers, those with wealth above w^E.

This equilibrium is similar to many practical schemes in which poor consumers receive free supplies while the rich do not, but this means-test equilibrium is not due to an equity concern. The public supplier selects among consumers with different wealth levels to participate in the private market. Wealthy consumers realize larger gains in trade in the private market, so they are rationed. The private market fully anticipates that poor consumers are unavailable, so even when cost decreases, the equilibrium price stops falling.

3.3. Characterization of a Continuum of Equilibria

In this subsection, we characterize all equilibria. We will show that in equilibrium the firm's quantity function is the monopoly quantity function for consumers with cost higher than a threshold, and a constant otherwise, but

this threshold must be higher than the one in Proposition 1. In equilibrium the public supplier must ration rich consumers, but may also ration some poor consumers.

To characterize the equilibrium rationing policies, we let the public supplier choose the net density of rationed consumers θf, and impose the requirement that $0 \leq \theta f \leq f$. The consumer welfare index (14) is linear in θf, and for each w its first-order derivative with respect to θf is

$$\frac{\partial V}{\partial \theta f} = \int_{\underline{c}}^{\widehat{c}(w)} \{U(w - \tau(\widehat{w}(c))) + 1 - U(w)\} \, g(c) \, \mathrm{d}c - 1. \qquad (16)$$

This expression measures the change in aggregate consumer utility at wealth level w. It is the expected incremental surplus from consumer with wealth w buying from the firm (the integral) less the unit incremental utility of consumption at zero cost. We establish a monotonicity in the supplier's preferences.

LEMMA 2: *The first-order derivative $\frac{\partial V}{\partial \theta f}$ in (16) is increasing in w. It is strictly increasing in $w \in [w_1, w_2]$ unless $\widehat{c}(w) = \underline{c}$ for each such w.*

Lemma 2 says that the public supplier favors rationing the consumer over supplying as the wealth level increases. This is a basic principle in our model. Prices in the private market depend on cost, so consumer (w, c) gets more surplus from a trade at price $\tau(\widehat{w}(c))$ as w increases: $U(w - \tau(\widehat{w}(c))) + 1 - U(w)$ is increasing in w. The public supplier's marginal utility from rationing, (16), is strictly increasing in w for consumers who buy from the firm. When consumers do not buy from the firm, there is no incremental surplus, so the integral in (16) is 0, and the derivative in (16) becomes -1, independent of w. Lemma 2 does not take into account the budget, the consideration of which is our next step.

Against a quantity function $\widehat{w}(c)$ (and the corresponding $\widehat{c}(w)$), the public supplier chooses θf to maximize (13) subject to the budget constraint (4). Using pointwise optimization, we consider the Lagrangean

$$\theta(w) f(w) \left[\begin{array}{l} \displaystyle\int_{\underline{c}}^{\widehat{c}(w)} \{U(w - \tau(\widehat{w}(c))) + 1\} \, g(c) \, \mathrm{d}c \\[2mm] + \displaystyle\int_{\widehat{c}(w)}^{\overline{c}} U(w) g(c) \, \mathrm{d}c \end{array} \right]$$

$$+ [1 - \theta(w)] f(w) [U(w) + 1] - \lambda[\gamma(1 - \theta(x)) f(x) - B],$$

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

$$\frac{\partial V}{\partial \theta f} + \lambda \gamma = \int_{\underline{c}}^{\widehat{c}(w)} \{U(w - \tau(\widehat{w}(c))) + 1 - U(w)\} \, g(c) \, \mathrm{d}c - 1 + \lambda \gamma. \quad (17)$$

From Lemma 2, the first-order derivative of the Lagrangean is strictly increasing in w whenever some consumers with wealth less than w purchase from the private market.

LEMMA 3: *In any equilibrium, the public supplier rations consumers with wealth above a threshold $\widetilde{w}$. That is, in an equilibrium there is $\widetilde{w} < \overline{w}$ such that $\theta(w) = 1$ for $w > \widetilde{w}$.*

Given the limited budget, some consumers must be rationed. There is always some scope for trade for rationed consumers at the private market because a consumer's willingness to pay is higher than the lowest cost. Now once there is some trade by rationed consumers, wealthier consumers get more incremental surplus. If it is optimal to ration a consumer ((17) positive at w), then it is also optimal to ration consumers richer than him ((17) strictly positive at $w' > w$).

Lemma 3 is consistent with the public supplier rationing some poor consumers in equilibrium. If the firm does not sell to very poor consumers, there is no incremental surplus for them. The first-order derivative (17) becomes $-1 + \lambda\gamma$, independent of w. In fact, when the value of (17) is 0, the public supplier is indifferent between rationing a consumer and supplying him. We can use this indifference to select various rationing rules to support equilibria.

LEMMA 4: *In any equilibrium, the private firm sets a constant price when cost falls below a threshold $\widetilde{c}$. That is, in an equilibrium there is $\widetilde{c} \leq \overline{c}$ such that $\widehat{w}(c)$ is constant for $c < \widetilde{c}$.*

Lemma 4 says that an equilibrium quantity function must become a constant when cost is sufficiently low. An equilibrium quantity function $\widehat{w}(c)$ is always increasing, and there must be some interaction between the two sectors because the budget satisfies $B > F(\widehat{w}^m(\underline{c}))\gamma$. If $\widehat{w}(c)$ is strictly increasing at low costs, similar to the solid line in Figure 1, the incremental surplus is higher for wealthy consumers. The public sector supplies poor consumers and rations wealthy consumers. When the poor consumers have been taken out of the market, it is no longer profit-maximizing for the firm to reduce price when cost becomes low. This is inconsistent with an equilibrium function being always strictly increasing at low costs.

The last two lemmas establish the form of an equilibrium, but do not pin down the exact strategies. Besides the equilibrium in Proposition 1, we can construct many equilibria exhibiting properties in Lemmas 3 and 4. Take the cost and wealth thresholds c^E and w^E in Proposition 1, and change the equilibrium rationing policy there to

$$\theta(w) = 1 \quad \text{for} \quad \underline{w} \leq w < \underline{w} + \epsilon$$

$$\theta(w) = 0 \quad \text{for} \quad \underline{w} + \epsilon < w < w^E + \delta$$
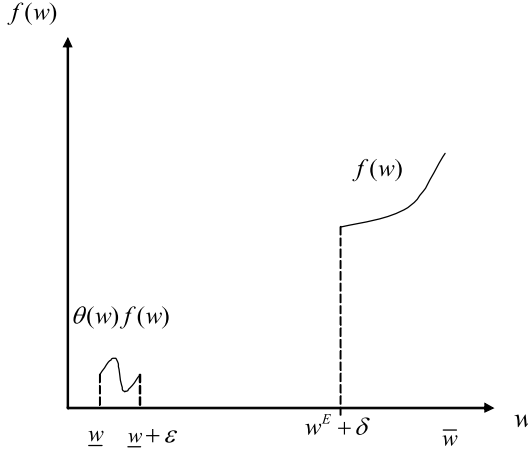
Figure 4: Equilibria in which some poor consumers are rationed.

$$\theta(w) = 1 \quad \text{for} \quad w^E + \delta < w \leq \overline{w},$$

where $\epsilon > 0$ and $\delta > 0$ are both small numbers.

In this rationing rule the supplier shifts some resources from those with wealth just above the lowest value $\underline{w}$ to those consumers with wealth just above w^E. Figure 4 shows the density of consumers available to the private firm in such an equilibrium.

Values of ϵ and δ can be so chosen that the new scheme satisfies the budget: $B = \gamma \left[F(w^E + \delta) - F(\underline{w} + \epsilon) \right]$. Against this rationing scheme, the private firm sets a quantity function equal to $\widehat{w}^m(c)$ for $c > c^E + \eta$ and $\widehat{w}^m(c^E + \eta)$ for $c < c^E + \eta$, where c^E is the cost threshold in Proposition 1, and $\eta > 0$ satisfies $\widehat{w}^m(c^E + \eta) = w^E + \delta$.

In this equilibrium, the public supplier gives the good to some consumers with wealth slightly higher than w^E, but rations consumers with wealth close to the lowest level. These rationed consumers have such low willingness to pay that the firm will not reduce price to sell to them even when cost is lowest. Furthermore, because consumers with wealth slightly higher than w^E are now supplied by the public, the private firm's price will not fall all the way to $\tau(w^E)$. In Appendix A, we provide a formal proof for this equilibrium.

Infinitely many equilibria can be constructed in a similar fashion. As long as the private firm does not find it profit-maximizing to reduce price to sell to consumers with low willingness to pay, a quantity function like the one in Figure 1 remains a best response. In all these equilibria the public supplier rations some consumers with low wealth, but must ration all consumers with wealth above a threshold.

The equilibrium in Proposition 1 is focal. This is the one that achieves the highest welfare index for the public supplier. This is because it has the widest range of price reduction as cost decreases. The equilibrium also allows the private firm to make the highest equilibrium profit. Any equilibrium different from the one in Proposition 1 would have fewer transactions in the private market.

PROPOSITION 2:   *The equilibrium in Proposition 1 achieves the highest equilibrium consumer utility, and the highest equilibrium profit for the private firm. In any other equilibrium, the public supplier sets $\theta(w) = 1$, for $w > \widetilde{w}^e$, where $\widetilde{w}^e > w^E$ (defined by $F(w^E)\gamma = B$ in Proposition 1) and the firm sets a price equal to $\tau(\widehat{w}^m(c))$ for $c > \widetilde{c}^e$, and a price equal to $\tau(\widetilde{w}^e)$ for $c < \widetilde{c}^e$, where $\widehat{w}^m(\widetilde{c}^e) = \widetilde{w}^e$ and $\widetilde{c}^e > c^E$ (defined by $w^E = \widehat{w}^m(c^E)$ in Proposition 1).*

How are consumers' utilities affected by the public supply and the price reaction in the private market? In Proposition 1, consumers with wealth above w^E are rationed, where $\gamma F(w^E) = B$, and the firm's equilibrium prices range between $\tau(\overline{w})$ and $\tau(w^E)$. Suppose that the budget increases by ΔB, then w^E will increase by Δw^E, where $\gamma F(w^E + \Delta w^E) = B + \Delta B$, so the minimum price in the private market becomes higher. In Figure 1, the new equilibrium is obtained by shifting the dotted horizontal line upward. An increase in the budget will be used in equilibrium to supply consumers with wealth just above w^E so available consumers in the private market are wealthier, and the firm reduces its price less when cost falls. Define Δc^E by $\widehat{w}^m(c^E + \Delta c^E) = w^E + \Delta w^E$. Consider consumers with wealth above $w^E + \Delta w^E$, those that remain rationed after the budget increase. They still are offered the monopoly prices when their costs are above $c^E + \Delta c^E$, so their utilities remain unchanged, but those with costs below $c^E + \Delta c^E$ face a higher price $\tau(w^E + \Delta w^E)$ although all of them still prefer to purchase. Wealthy and low-cost consumers are hurt by the budget increase, while more poor consumers benefit.

COROLLARY 1:   *Under rationing based on wealth, private market equilibrium prices are higher when the public supplier's budget increases. Consumers who remain rationed after the budget increase face a strictly higher price when their costs are low.*

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

Next, we can extend our model to the case of a perfectly competitive private sector. Here, the price in the private sector is marginal cost: $\tau(c) = c$. Given this pricing function, the corresponding quantity function $\widehat{w}(c)$ is implicitly defined by $U(\widehat{w} - c) + 1 = U(\widehat{w})$. Lemma 2 can be applied to this quantity function. Because the perfectly competitive quantity function is strictly increasing, the derivative (17) is strictly increasing for all values of w. Lemma 3 continues to hold. In sum we have the following.

COROLLARY 2:   *If the private market is perfectly competitive so that prices are equal to marginal costs, the public sector uses the entire budget on consumers with low wealth levels:* $\theta(w) = 0$, *for* $w < w^E$, *and* $\theta(w) = 1$, *for* $w > w^E$ *where* $F(w^E)\gamma = B$.

4. Equilibrium Rationing and Prices when Rationing is Based on Wealth and Cost

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

$$\int_w^{\overline{w}} \phi(x, c) f(x)\mathrm{d}x\, [\tau(w) - c]. \tag{18}$$

We use the same notation and let $\widehat{w}(c)$ maximize profit (18).

Consider a rationing function $\phi : [\underline{w}, \overline{w}] \times [\underline{c}, \overline{c}] \to [0, 1]$ and a quantity function $\widehat{w} : [\underline{c}, \overline{c}] \to [\underline{w}, \overline{w}]$. Consumer (w, c) buys from the private firm if and only if $U(w - \tau(\widehat{w}(c))) + 1 \geq U(w)$. Therefore, when the supplier rations consumer (w, c), that consumer obtains a utility $\max[U(w - \tau(\widehat{w}(c))) + 1,\ U(w)]$ from the private sector. Aggregate consumer utility from a policy ϕ is

$$\int_{\underline{c}}^{\overline{c}} \int_{\underline{w}}^{\overline{w}} \{\phi(w, c) \max[U(w - \tau(\widehat{w}(c))) + 1,\ U(w)]$$
$$+\ [1 - \phi(w, c)][U(w) + 1]\} f(w)g(c)\, \mathrm{d}w\, \mathrm{d}c. \tag{19}$$

Our next Proposition reports that in the unique equilibrium the public supplier's allocation rule depends only on consumers' cost level. Consumers are supplied the good if and only if their costs are lower than a threshold, irrespective of their wealth levels. The rationing rule based on w and c is the same as the optimal allocation without a private sector.

PROPOSITION 3:   *If the public supplier rations consumers based on wealth and cost information, the equilibrium rationing function is identical to the optimal rationing function when the private sector is inactive. In the equilibrium, consumer*

(w, c) is rationed when his cost is above a threshold $c^s > \underline{c}$, and supplied when his cost is below c^s. That is, $\phi(w, c) = 1$ for $c > c^s$, any w, and $\phi(w, c) = 0$ otherwise, where $\int_{\underline{c}}^{c^s} c \, d\, G(c) = B$. The private firm chooses the monopoly quantity $\widehat{w}^m(c)$ for $c > c^s$, and $\widehat{w}(c) = \overline{w}$ for $c < c^s$.

When the supplier observes wealth and cost information, a standard cost-effectiveness principle applies. If the cost is too high relative to the benefit, the consumer should not be given the good. Now the availability of the private supply to high-cost consumers does not alter this principle. Next, consider low-cost consumers. The cost-effectiveness principle points against rationing. However, some wealthy and low-cost consumers may obtain a higher surplus from the private sector if the price is low.

Might the public supplier want to ration them, and allow them to get the surplus from the private market? Proposition 3 says that this cannot happen in equilibrium. Consider the marginal consumer (w, c); he pays a price $\tau(w)$ in the private market and obtains a zero incremental surplus. Now if his cost is low, the public supplier prefers to allocate the good to him, giving him a positive incremental surplus. By continuity, the supplier also assigns the good to those with wealth slightly above w, eliminating these consumers from the private market. Given that rationed consumers are richer, the best response by the private firm is to raise the price. This unravelling continues until the private firm raises the price to $\tau(\overline{w})$. In an equilibrium, it is as if low-cost consumers did not have a private option, so the supplier provides the good to them.

Proposition 3 is consistent with practical cost-effectiveness policies. We have used a standardized benefit of one unit of utility increment from the consumption of the good. Therefore, Proposition 3 should be understood to say that in equilibrium only consumers with sufficiently low costs per unit of benefit will be supplied. This is indeed what most public programs aim to achieve.

In contrast to the regime when rationing is based on wealth information only, here consumers are not hurt by public supply. An increase in the budget, say by ΔB, will raise the value of c^s, say by Δc^s, so more consumers benefit from public supply. The price schedule will not change. Those consumers who remain rationed after the budget increase face the same monopoly price schedule. Let $\widehat{w}^m(c^s + \Delta c^s) = w^s + \Delta w^s$. The prices for consumers with cost above $c^s + \Delta c^s$ now range between $\tau(w^s + \Delta w^s)$ to $\tau(\overline{w})$, the same as before the budget increase.

COROLLARY 3: *Under rationing based on wealth and cost, rationed consumers face the same equilibrium prices when the budget increases.*

Clearly, Proposition 3 applies directly to Cournot competition in the private market. Equilibrium rationing when the private market is perfectly competitive is a little different. Here, some low-cost and wealthy consumers
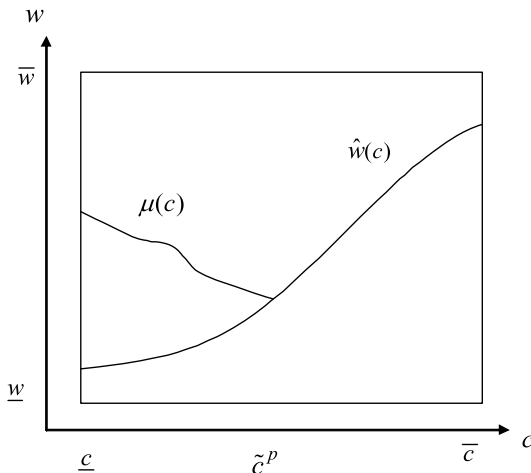
Figure 5: Optimal rationing in a competitive private market.

may be rationed. When the private market is competitive, prices are given by marginal costs, $\tau(c) = c$. The consumer welfare index is

$$\int_{\underline{c}}^{\overline{c}} \int_{\underline{w}}^{\overline{w}} \{\phi(w, c) \max[U(w - c) + 1, U(w)]$$

$$+ [1 - \phi(w, c)][U(w) + 1]\} f(w)g(c) \, \mathrm{d}w \, \mathrm{d}c. \qquad (20)$$

COROLLARY 4: *Suppose the private market is competitive so that prices are equal to marginal costs. The public sector rations all consumers with cost above a threshold. For those consumers with costs below the threshold, the public sector supplies a consumer if and only if his wealth is below a value that is determined by his cost. That is, $\phi(w, c) = 1$ for $c > \tilde{c}^p$, some $\tilde{c}^p > \underline{c}$ and any w; $\phi(w, c) = 1$ for $c < \tilde{c}^p$ and $w > \mu(c)$ where the function μ is implicitly defined by $U(\mu - c) - U(\mu) + \lambda c = 0$ for a constant $\lambda > 0$; otherwise, $\phi(w, c) = 0$.*

Corollary 4 is illustrated in Figure 5. The upward sloping line $\widehat{w}(c)$ is the marginal consumer given marginal cost pricing. The line above $\widehat{w}(c)$ is the function μ defined in the Corollary.[7] Consumers with cost $c < \tilde{c}^p$ and wealth between $\widehat{w}(c)$ and $\mu(c)$ strictly prefer to purchase the good at cost, yet the supplier will assign the good to them for free. Wealthy consumers, those with $w > \mu(c)$, are rationed even when their costs are below $\tilde{c}^p$.

For consumers with costs higher than $\tilde{c}^p$, it is not cost effective to supply them. The logic in the first part of the proof of Proposition 3 applies. For those with costs lower than $\tilde{c}^p$, the logic in Proposition 3 applies with

---

[7] By definition, $U(\widehat{w}(c) - c) + 1 = U(\widehat{w}(c))$. Now because $c < \tilde{c}^p$, we have $1 > \lambda c$. It follows that the value of μ that satisfies $U(\mu - c) + 1 = U(\mu) + (1 - \lambda c)$ must be greater than $\widehat{w}(c)$. We have drawn μ to be negatively sloped, but it does not have to be.

one modification. Cost consideration alone warrants the allocation. Nevertheless, wealthy consumers obtain higher incremental surplus from the private market. The condition in Corollary 4 can be rewritten as

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

The left-hand side expression is the utility of a consumer with wealth μ buying from the competitive market; the right-hand side expression is the social net benefit, where λ is the multiplier of the budget constraint and λc measures the utility equivalent of cost c. If $w > \mu(c)$, the consumer gets more utility from the private market than the social net benefit from the good, so will be rationed. In contrast to Proposition 3, there is no unravelling of equilibrium price best responses in the private market.

4.1. Equilibrium Consumer Utilities in Rationing Regimes

We now compare equilibrium consumer utilities across the two rationing regimes in terms of the budget B. For this comparison, we do not need any restriction on the size of the budget,[8] so we let $0 < B < \gamma$. For rationing based on wealth, we take the equilibrium in Proposition 1, and let $V^w(B)$ denote consumers' equilibrium utility. For rationing based on wealth and cost, we take the equilibrium in Proposition 3, and let $V^{wc}(B)$ denote consumers' equilibrium utility.

PROPOSITION 4: *Equilibrium consumer utility under rationing based on wealth is strictly less than equilibrium consumer utility under rationing based on wealth and cost.*

The following are expressions for $V^w(B)$ and $V^{wc}(B)$:

$$V^w(B) = \int_{\underline{w}}^{\overline{w}} U(w) f(w) \, dw + F(w^E)$$

$$+ \int_{w^E}^{\overline{w}} \left\{ \begin{array}{l} \int_{\underline{c}}^{c^E} [U(w - \tau(w^E)) + 1 - U(w)] g(c) \, dc \\ \\ + \int_{c^E}^{\widehat{c}^m(w)} [U(w - \tau(\widehat{w}^m(c))) + 1 - U(w)] g(c) \, dc \end{array} \right\} f(w) \, dw,$$

$$V^{wc}(B) = \int_{\underline{w}}^{\overline{w}} U(w) f(w) \, dw + G(c^s)$$

$$+ \int_{c^s}^{\overline{c}} \left\{ \int_{\widehat{w}^m(c)}^{\overline{w}} [U(w - \tau(\widehat{w}^m(c))) + 1 - U(w)] f(w) \, dw \right\} g(c) \, dc.$$

Both V^w and V^{wc} contain the base utility $\int_{\underline{w}}^{\overline{w}} U(w) f(w) \, dw$. In V^w, the consumption utility from public supply is $F(w^E)$; in V^{wc}, it is $G(c^s)$. Each consumer gets one unit of utility from consumption, and $F(w^E)$ and $G(c^s)$ are

---

the masses of supplied consumers in the two equilibria. Under the budget B, we have $F(w^E)\gamma = \int_{\underline{c}}^{c^s} c \, dG(c) = B$. For any B between 0 and γ, $c^s < \overline{c}$, so $\int_{\underline{c}}^{c^s} c \, dG(c) < \gamma \, G(c^s)$, and $G(c^s) > F(w^E)$. Cost-effectiveness allows the public supplier to provide the good to more consumers.

The remaining, double integrals in V^w and V^{wc} are consumers' incremental surplus in the market in each equilibrium. Consider a consumer (w, c). If he participates in the market under wealth-based rationing, he pays the price $\tau(w^E)$ or $\tau(\widehat{w}^m(c))$. If he participates in the market under wealth-cost rationing, he pays the price $\tau(\widehat{w}^m(c))$. If $c < c^E$, $\tau(w^E) > \tau(\widehat{w}^m(c))$, the consumer pays a higher price under wealth rationing. If $c > c^E$, the consumer pays the same price. Conditional on participating in the market, a consumer's incremental surplus is higher under wealth-cost rationing. Nevertheless, the sets of consumers who trade in the private market may be different in the two equilibria.

An increase in the budget has two effects on the values of $V^w(B)$ and $V^{wc}(B)$. The first is that public supply increases, so c^s and w^E increase. This is a first-order effect. Also, this effect is always stronger under wealth-cost rationing: $G(c^s) > F(w^E)$. The second is that incremental surplus in the private market is affected. The double integrals in $V^w(B)$ and $V^{wc}(B)$ depend on c^s and w^E. Nevertheless, consumers are making optimal decisions, so the impact of B on the incremental is of second order.

Clearly, when the budget is 0, equilibria across the two regimes are the same because the public supplier must ration all consumers. When the budget is γ (the expected cost), all consumers can be supplied in both regimes. Hence, we can conclude that $V^w(0) = V^{wc}(0)$, and $V^w(\gamma) = V^{wc}(\gamma)$. Now when B starts at 0, because $G(c^s) > F(w^E)$, we must have $V^{wc}(B)$ increasing faster than $V^w(B)$, so $V^{wc}(B) > V^w(B)$, which is then maintained through the range $0 < B < \gamma$.

In our games, the public supplier and the firm move simultaneously. Generally, more information need not benefit a player in games with simultaneous moves. Nevertheless, Proposition 4 says that rationing based on costs allows a much better use of limited resources than rationing based on wealth. Providing the good to those with a low cost-benefit ratio generates the largest benefit for a given budget. Cream-skimming in the private sector does not change the superiority of cost-effectiveness over means-testing rationing. Proposition 4 yields the policy implication that obtaining cost information is worthwhile.

5. Concluding Remarks

We have introduced a framework for studying strategic interactions between public and private sectors. In our model, the public sector uses nonprice rationing and the private sector uses a pricing rule. We derive equilibria when rationing is based on consumers' wealth or both wealth and cost

information. Equilibria do look like common rationing schemes: means-testing and cost effectiveness schemes. Prices in the private sector are high because rich or high-cost consumers seek services there.

Our model can be adapted for studying monetary subsidies and their effects on pricing in the private sector. It can also be used for studying quality differences between the public and private sectors. Appendix B reviews some of the possible issues for more general model specifications. While we believe that our results are robust against small changes in model specifications, extending the model to include the general specifications may be worthwhile.

Finally, we have assumed a fixed budget for the public supplier. Extending the model to an endogenous budget is straightforward. Given a budget, our model yields equilibrium outcomes in Propositions 1 and 3. Moreover, the equilibrium values for the consumer welfare index can be obtained. Once the social cost for the budget is specified, the optimal level for the budget can be studied.

Appendix A

Proof of Lemma 1:   For $c_1 < c_2$, let $w_1 \in \widehat{w}(c_1)$ and $w_2 \in \widehat{w}(c_2)$. Because the profit function $\pi(w; c, \theta)$ in (7) is strictly decreasing in c, we have $\widehat{\pi}(c_1) \geq \pi(w_2; c_1, \theta) > \pi(w_2; c_2, \theta) = \widehat{\pi}(c_2)$. Hence, the maximum profit function $\widehat{\pi}(c)$ is strictly decreasing in c.

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

$$\int_{w_1}^{\overline{w}} \theta(w) f(w) \, \mathrm{d}w \, [\tau(w_1) - c_1] \geq \int_{w_2}^{\overline{w}} \theta(w) f(w) \, \mathrm{d}w \, [\tau(w_2) - c_1]$$

$$\int_{w_2}^{\overline{w}} \theta(w) f(w) \, \mathrm{d}w \, [\tau(w_2) - c_2] \geq \int_{w_1}^{\overline{w}} \theta(w) f(w) \, \mathrm{d}w \, [\tau(w_1) - c_2].$$

Adding these two inequalities yields

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

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

Proof of Proposition 1:   In the discussion following the proposition, we already show that the private firm's strategy is optimal against the public supplier's rationing policy. It remains to show that the public supplier's policy is optimal, and this is equivalent to showing that the public supplier's strategy in the proposition solves the maximization of (15) subject to the budget constraint (4). We use pointwise optimization to find the solution of this constrained

optimization problem. Furthermore, for convenience, we let the choice be the function θf, and impose the boundary conditions, $0 \leq \theta f \leq f$.

For $w < w^E$, the Lagrangean is

$$L \equiv U(w) + (1 - \theta(w))\, f(w) + \lambda[B - \gamma(1 - \theta(w))\, f(w)].$$

For $w > w^E$, it is

$$L \equiv U(w) + (1 - \theta(w))\, f(w) + \lambda[B - \gamma(1 - \theta(w))\, f(w)]$$

$$+ \left\{ \begin{array}{l} \displaystyle\int_{\underline{c}}^{c^E} [U(w - \tau(w^E)) + 1 - U(w)]\, g(c)\, \mathrm{d}c \\[4mm] \displaystyle + \int_{c^E}^{\widehat{c}^m(w)} [U(w - \tau(\widehat{w}^m(c))) + 1 - U(w)]\, g(c)\, \mathrm{d}c \end{array} \right\} \theta(w)\, f(w).$$

For $w < w^E$, the first-order derivative of L with respect to θf is

$$\frac{\partial L}{\partial \theta f} = -1 + \lambda\gamma. \tag{A1}$$

For $w > w^E$, the first-order derivative of L with respect to θf is

$$\frac{\partial L}{\partial \theta f} = -1 + \lambda\gamma$$

$$+ \left\{ \begin{array}{l} \displaystyle\int_{\underline{c}}^{c^E} [U(w - \tau(w^E)) + 1 - U(w)]\, g(c)\, \mathrm{d}c \\[4mm] \displaystyle + \int_{c^E}^{\widehat{c}^m(w)} [U(w - \tau(\widehat{w}^m(c))) + 1 - U(w)]\, g(c)\, \mathrm{d}c \end{array} \right\}. \tag{A2}$$

We set λ at $1/\gamma$, and verify that the rationing policy in the proposition together with this λ satisfy conditions for the solution of the constrained optimization problem. Indeed, at $\lambda = 1/\gamma$, the value of (A1) is 0, so it is optimal to set $\theta(w)$ to 0, and the value of (A2) is strictly positive, so it is optimal to set $\theta(w)$ to 1. This is the rationing policy in the proposition. Furthermore, by the definition of w^E, the budget constraint is satisfied. Finally, if λ were different from $1/\gamma$, the budget constraint is either slack or violated. There is no other solution for which $\lambda \neq 1/\gamma$. ∎

Proof of Lemma 2: Consider w_1 and w_2 with $w_1 < w_2$. Evaluating (16) at w_1 and w_2 and then taking the difference, we have

$$\left. \frac{\partial V}{\partial \theta f} \right|_{w=w_2} - \left. \frac{\partial V}{\partial \theta f} \right|_{w=w_1}$$

$$= \int_{\underline{c}}^{\widehat{c}(w_1)} \{ [U(w_2 - \tau(\widehat{w}(c)))$$

$$- U(w_1 - \tau(\widehat{w}(c)))] - [U(w_2) - U(w_1)] \} g(c) \, dc$$

$$+ \int_{\widehat{c}(w_1)}^{\widehat{c}(w_2)} \{ U(w_2 - \tau(\widehat{w}(c))) + 1 - U(w_2) \} g(c) \, dc \geq 0 \qquad \text{(A3)}$$

The inequality in (A3) follows from the concavity of U, and $\widehat{c}$ being increasing whenever $\widehat{c}(w) \neq \underline{c}$. Finally, (A3) is zero if and only if $\widehat{c}(w_1) = \widehat{c}(w_2) = \underline{c}$. ∎

Proof of Lemma 3: Let $w^* = \inf\{w : \theta(w) > 0\}$. Due to the limited budget, the public supplier must leave some consumers to the private sector, so $w^* < \overline{w}$. Because $\tau(\underline{w}) > \underline{c}$, the firm must sell to some consumer, say, $\widetilde{w} > w^*$, so $c(\widetilde{w})$ must be higher than $\underline{c}$. By Lemma 2, the first-order derivative of the Lagrangean with respect to θf must be strictly increasing in w for $w > \widetilde{w}$. Since $\theta(\widetilde{w}) > 0$, the first-order derivative (17) must be nonnegative at $\widetilde{w}$, and for any $w > \widetilde{w}$, the value of (17) must be strictly positive, and $\theta(w) = 1$. ∎

Proof of Lemma 4: Let $\widehat{w}(c)$ be an equilibrium quantity function. This is increasing by Lemma 1. First, if $\widetilde{c} = \underline{c}$, the Lemma is vacuously true, so we let $\widetilde{c} > \underline{c}$. Suppose that the Lemma is false. That is, suppose that for some $\widetilde{c} \leq \overline{c}$, $\widehat{w}(c)$ is strictly increasing for all $\underline{c} < c < \widetilde{c}$. The firm sells to $\widehat{w}(c)$, so all consumers within $[\widehat{w}(\underline{c}), \widehat{w}(\widetilde{c})]$ must be rationed with positive probabilities. This means that the first-order derivative (17) of the Lagrangean at $\underline{c}$ must be nonnegative. Furthermore, $\widehat{w}(c)$ strictly increasing implies that $\widehat{c}(w) > \underline{c}$ for all w in $[\widehat{w}(\underline{c}), \widehat{w}(\widetilde{c})]$. By Lemma 2, the first-order derivative (17) must be strictly increasing at any w in $[\widehat{w}(\underline{c}), \widehat{w}(\widetilde{c})]$, and therefore strictly positive for all $w > \widehat{w}(\underline{c})$. The public supplier rations all consumers with wealth above $\widehat{w}(\underline{c})$.

The budget is then allocated to all consumers with $w < \widehat{w}(\underline{c})$. The budget satisfies $B > F(\widehat{w}^m(\underline{c}))\gamma$, so $\widehat{w}(\underline{c}) > \widehat{w}^m(\underline{c})$. For $\epsilon > 0$ and arbitrarily small, we have $\widehat{w}(\underline{c} + \epsilon) > \widehat{w}^m(\underline{c} + \epsilon)$ because $\widehat{w}(\underline{c}) > \widehat{w}^m(\underline{c})$. The derivative of the profit function $\int_w^{\overline{w}} f(x) \, dx \ \ [\tau(w) - c - \epsilon]$ at $w = \widehat{w}(\underline{c} + \epsilon)$ is negative by (10), so setting w below $\widehat{w}(\underline{c} + \epsilon)$ is feasible and yields a higher profit. This contradicts the assumption that $\widehat{w}(\underline{c} + \epsilon)$ maximizes the firm's profit. We conclude that there must be $\widetilde{c} \leq \overline{c}$ such that $\widehat{w}(c)$ is constant for $c < \widetilde{c}$. ∎

Proof Rationing Rule in Figure 4 Constitutes an Equilibrium: Given the rationing
rule, for $c > c^E + \eta$, the profit-maximizing price is $\tau(\widehat{w}^m(c))$, and profit is

$$\int_{\widehat{w}^m(c)}^{\overline{w}} f(x) \, dx \, [\tau(\widehat{w}^m(c)) - c].$$

There are no consumers with wealth between $\underline{w} + \epsilon$ and $w^E + \delta$, so price
remains at $\tau(w^E + \delta)$ when cost reduces from $c^E + \eta$ to $\tau(\underline{w} + \epsilon)$. For costs
below $\tau(\underline{w} + \epsilon)$, the firm must drop the price below $\tau(\underline{w} + \epsilon)$ to sell to con-
sumers with w between $\underline{w}$ and $\underline{w} + \epsilon$. For $w < \underline{w} + \epsilon$, the profit is

$$\int_{\underline{w}}^{\underline{w}+\epsilon} f(x) \, dx \, [\tau(w) - c] + \int_{w^E+\delta}^{\overline{w}} f(x) \, dx \, [\tau(w) - c].$$

If the firm does not sell to consumers with wealth between $\underline{w}$ and $\underline{w} + \epsilon$, it
sets the price at $\tau(w^E + \delta)$, and profit is

$$\int_{w^E+\delta}^{\overline{w}} f(x) \, dx \, [\tau(w^E + \delta) - c].$$

Setting the price at $\tau(w^E + \delta)$ yields higher profit because $\tau(w^E + \delta) > \tau(w)$
when $\underline{w} \le w \le \underline{w} + \epsilon$ and ϵ sufficiently small. We conclude that the price
remains at $\tau(w^E + \delta)$ when cost falls below $c^E + \eta$.

The rest of the proof is similar to that of Proposition 1. Given the firm's
strategy, at $w < w^E + \delta$, the Lagrangean of the public supplier's optimization
has a zero derivative, while at $w > w^E + \delta$ it is strictly positive. It is optimal
for the public supplier to ration all consumers with $w > w^E + \delta$. For $w <
w^E + \delta$, any rationing policy is optimal, so we choose the one in Figure 4. We
can choose ϵ from an open set to construct a continuum of equilibria. ∎

Proof of Proposition 2: In any equilibrium, the budget constraint $\gamma \int_{\underline{w}}^{\overline{w}} (1 -
\theta(x)) f(x) \, dx \le B$ must hold as an equality. The equilibrium in Proposition
1 supplies those with wealth between $\underline{w}$ and w^E, where $F(w^E)\gamma = B$, so
$\theta(w) = 0$ for $w < w^E$ and $\theta(w) = 1$ for $w > w^E$. Consider any other equi-
librium. Here, the supplier must ration some consumers with wealth be-
low w^E because the budget constraint must hold. Hence, for this equilib-
rium the threshold $\widetilde{w}^e$ at which $\theta(w) = 1$ for $w > \widetilde{w}^e$ must be strictly higher
than w^E.

Let $\widetilde{c}^e$ be defined by $\widehat{w}^m(\widetilde{c}^e) = \widetilde{w}^e$. Clearly, $\widetilde{c}^e > c^E$ because $\widehat{w}^m$ is strictly
increasing and $\widetilde{w}^e > w^E$. We now show that the firm's equilibrium price is
$\tau(\widehat{w}^m(\widetilde{c}^e))$ for $c < \widetilde{c}^e$. At $w < \widetilde{w}^e$, the first-order derivative of the Lagrangean
(17) must be nonnegative; if that derivative was negative, then $\theta(w) = 0$,
which would violate the budget constraint. Because $\theta(w) = 1$ for $w > \widetilde{w}^e$, the
value of (17) is positive for all $w > \widetilde{w}^e$. Therefore, by Lemma 2, for $w < \widetilde{w}^e$
the value of (17) must be exactly zero and $\widehat{c}(w) = \underline{c}$ for $w < \widetilde{w}^e$. Because
$\theta(w) = 1$ for $w > \widetilde{w}^e$ the equilibrium quantity must be $\widehat{w}^m(c)$ for $c > \widetilde{c}^e$, and
remains constant at $\widehat{w}^m(\widetilde{c}^e)$ for $c < \widetilde{c}^e$.

From $c^E < \tilde{c}^e$ and $\tilde{w}^e > w^E$, by comparing the values of (13) across the equilibrium in Proposition 1 and the alternative, we conclude that the supplier's payoff is higher in the equilibrium in Proposition 1.

Finally, consider the equilibrium price function in Proposition 1 and any other equilibrium price function. The cost threshold is c^E in Proposition 1, and $\tilde{c}^e$ in the other, where the prices remain constant at $\tau(\widehat{w}^m(c^E))$ and $\tau(\widehat{w}^m(\tilde{c}^e))$ respectively when cost falls below these thresholds. The private firm sets the same price and sells to the same set of consumers when $c > \tilde{c}^e$, making the same profit in the two equilibria. Now for $c < \tilde{c}^e$, in the equilibrium in Proposition 1, the firm could have set a price at $\tau(\widehat{w}^m(\tilde{c}^e))$ and sell to fewer consumers but has chosen to set a lower price at $\tau(\widehat{w}^m(c))$ or $\tau(\widehat{w}^m(c^E))$. Therefore, the firm makes more profit at $c < \tilde{c}^e$ in the equilibrium in Proposition 1. We conclude that the private firm makes the largest equilibrium expected profit in the equilibrium in Proposition 1. ∎

Proof of Proposition 3: In an equilibrium, the public supplier chooses ϕ to maximize (19) subject to the budget constraint (6), given a quantity function $\widehat{w}(c)$. Consider pointwise maximization at each (w, c). The Lagrangean is

$$\{\phi(w, c) \max [U(w - \tau(\widehat{w}(c))) + 1, U(w)]$$
$$+ [1 - \phi(w, c)] [U(w) + 1]\} f(w) g(c)$$
$$+ \lambda [B - c(1 - \phi(w, c)) f(w) g(c)]. \tag{A4}$$

The first-order derivative of (A4) with respect to ϕfg is

$$\max [U(w - \tau(\widehat{w}(c))) + 1, U(w)] - [U(w) + 1] + \lambda c, \tag{A5}$$

where $\lambda > 0$ is the multiplier.

Given a quantity function, if consumer (w, c) prefers to purchase from the private firm, the expression in (A5) becomes

$$U(w - \tau(\widehat{w}(c)) - U(w) + \lambda c. \tag{A6}$$

Otherwise, the expression in (A5) becomes

$$-1 + \lambda c. \tag{A7}$$

When $U(w - \tau(\widehat{w}(c))) + 1 > U(w)$, the value in (A6) is larger than (A7).

Consider all consumers (w, c) who do not purchase from the private firm. Now the first-order derivative is given by (A7), which is strictly increasing in c. Furthermore, if $-1 + \lambda c > 0$, then the expression in (A6) is also strictly positive.

Now we claim that in an equilibrium expression (A7) must not be always strictly positive. Suppose not, then $-1 + \lambda c > 0$ for all c, and therefore, the first-order derivative (A5) is always strictly positive. The supplier rations all consumers so that $\phi(w, c) = 1$ for all w and c. This implies that the supplier

does not use its budget, and this cannot be optimal. We conclude that there must exist $\widetilde{c} > \underline{c}$ such that (A7) vanishes at $c = \widetilde{c}$.

Now, whenever $c > \widetilde{c}$, $-1 + \lambda c > 0$ so the first-order derivative (A5) is strictly positive. We conclude that $\phi(w, c) = 1$, all $c > \widetilde{c}$ and any w.

Next, consider $c < \widetilde{c}$. The value of (A7) is strictly negative for $c < \widetilde{c}$. For $w < \widehat{w}(c)$, $U(w - \tau(\widehat{w}(c))) + 1 < U(w)$ so that (A5) takes the value of (A7) which is strictly negative. We conclude that $\phi(w, c) = 0$.

In equilibrium the budget must be exhausted. From the above properties of ϕ, we must have $\widetilde{c} = c^s$ where $B = \int_{\underline{c}}^{c^s} g(c)\,dc$. Furthermore, we have $\lambda = 1/c^s$ because at $c = c^s$, (A7) vanishes.

We now show that the firm's quantity function in the Proposition is the unique equilibrium strategy. Clearly, for $c > c^s$, the firm choosing $\widehat{w}^m(c)$ is the unique best response. The proof is complete if for $c < c^s$ the firm's unique equilibrium strategy is $\widehat{w}(c) = \overline{w}$. Suppose not; that is, suppose that for $c < c^s$, $\widehat{w}(c) < \overline{w}$. Now at $w = \widehat{w}(c)$, $U(w - \tau(\widehat{w}(c))) + 1 = U(w)$, so that the derivative (A5) is negative. For $\epsilon > 0$ and sufficiently small, $U(w + \epsilon - \tau(\widehat{w}(c))) - U(w + \epsilon) + \lambda c < 0$, so that the derivative (A5) remains negative, and $\phi(w + \epsilon, c) = 0$. Now given that at c, consumers with wealth between $w = \widehat{w}(c)$ and $w = \widehat{w}(c) + \epsilon$ are supplied by the public sector, the private firm will raise the price from $\tau(\widehat{w}(c))$, or equivalently raise the equilibrium quantity function from $\widehat{w}(c)$. This contradicts the assumption that $\widehat{w}(c) < \overline{w}$ is an equilibrium quantity function. We conclude that for $c < c^s$, $\widehat{w}(c) = \overline{w}$. ∎

Proof of Corollary 4: First, set $\tau(\widehat{w}(c))$ to c in the proof of Proposition 3. It follows that $\phi(w, c) = 1$ for $c > \widetilde{c}^p$, some $\widetilde{c}^p > \underline{c}$. Now consider $c < \widetilde{c}^p$. For those consumers (w, c) who do not purchase from the private sector, $\phi(w, c) = 0$.

Consider consumer (w, c), $c < \widetilde{c}^p$, and $U(w - c) + 1 > U(w)$, so that this consumer purchases from the private sector at cost c. Setting $\tau(\widehat{w}(c))$ to c in (A6) yields the first-order derivative of the Lagrangean with respect to ϕ,

$$U(w - c) - U(w) + \lambda c, \tag{A8}$$

where $\lambda > 0$ is the multiplier for the budget constraint. Expression (A8) is negative at $c < \widetilde{c}^p$ and $w = \widehat{w}(c)$. Hence, if there exists $w' > \widehat{w}(c)$ such that $U(w' - c) - U(w') + \lambda c > 0$, then $\phi(w', c) = 1$. The function $\mu(c)$ in the corollary is implicitly defined by setting the first-order derivative (A8) to 0. ∎

Proof of Proposition 4: The function $V^w(B)$ is the maximized consumer utility in the equilibrium in Proposition 1. By the Envelope Theorem, the derivative of V^w is the partial derivative of the corresponding Lagrangean with respect to B, which is the multiplier. From the proof of Proposition 1, this multiplier

is $1/\gamma$:

$$\frac{\mathrm{d}V^w(B)}{\mathrm{d}B} = \frac{1}{\gamma} \qquad (A9)$$

which is a constant, so that $V^w(B)$ is linear in B.

Similarly, from the Envelope Theorem:

$$\frac{\mathrm{d}V^{wc}(B)}{\mathrm{d}B} = \frac{1}{c^s}, \qquad (A10)$$

where c^s, given by $\int_{\underline{c}}^{c^s} c \, \mathrm{d}G(c) = B$, is the reciprocal of the multiplier of the Lagrangean in the proof of Proposition 3.

Now, c^s is a strictly increasing function of B. Therefore, the derivative of V^{wc} in (A10) is strictly decreasing in B, and V^{wc} strictly concave. Together with $V^w(0) = V^{wc}(0)$, and $V^w(\gamma) = V^{wc}(\gamma)$, we conclude that $V^w(B) < V^{wc}(B)$ for $0 < B < \gamma$. ■

Appendix B

In this appendix, we discuss the general, nonseparable utility function, the situation where benefits may depend on costs, and the possible correlation between wealth and costs.

B.1. General Utility Function

If we use a general utility specification, the willingness-to-pay function is implicitly defined by $U(w - \tau, 1) = U(w, 0)$. From total differentiation, we have

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

The willingness-to-pay function is increasing if the goods are complements ($\frac{\partial U}{\partial w}(w, 1) > \frac{\partial U}{\partial w}(w, 0)$), but may still be increasing even when they are substitutes ($\frac{\partial U}{\partial w}(w, 1) < \frac{\partial U}{\partial w}(w, 0)$).

Consider first optimal rationing based on wealth. The welfare index is $\int_{\underline{w}}^{\overline{w}} [1 - \theta(w)] U(w, 1) \, \mathrm{d}F + \int_{\underline{w}}^{\overline{w}} \theta(w) U(w, 0) \, \mathrm{d}F$. When the private market is inactive, the optimal rationing rule maximizes the welfare index subject to the budget constraint. The Lagrangean is $[1 - \theta] U(w, 1) + \theta U(w, 0) + \lambda[B - (1 - \theta)\gamma]$, and the first-order derivative with respect to θ is $-U(w, 1) + U(w, 0) + \lambda\gamma$. This first-order derivative is increasing in w if and only if wealth and the good are substitutes. In this case, the public sector supplies the good to consumers with low wealth, so that $\theta(w) = 0$ if w is below a threshold, and $\theta(w) = 1$ otherwise.

When the willingness-to-pay function is increasing and when the optimal rationing rule favors consumers with low wealth levels, our arguments in Section 3 apply. Results in Section 3 will have to be modified if the two goods are strong substitutes because then the public supplier may ration poor consumers.

Next, consider rationing based on wealth and cost. We can again derive the optimal rationing scheme when the private sector is inactive. The first-order derivative with respect to the rationing probability ϕ is $-U(w, 1) + U(w, 0) + \lambda c$, which is strictly increasing in c, and strictly increasing in w if and only if wealth and the good are substitutes.

Define (w^*, c^*) by $-U(w^*, 1) + U(w^*, 0) + \lambda c^* = 0$. If wealth and the good are substitutes, the optimal rationing rule sets $\phi(w, c) = 0$ if and only if $w \leq w^*$ and $c \leq c^*$. If they are complements, the optimal rationing rule sets $\phi(w, c) = 0$ if and only if $w \geq w^*$ and $c \leq c^*$. In either case, the optimal cost-effectiveness rule assigns the good to low-cost consumers.

In sum, the separable utility function is convenient. We have abstracted from secondary wealth effects due to the consumption of the good. In any case, when the consumption of the good results in small changes in marginal utilities of income, our results may continue to hold.

B.2. Benefits Dependent on Costs

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

Now suppose that the supplier has both wealth and cost information. The standard cost-effectiveness comparison will be modified as follows. If the private market is inactive, the comparison is between $\beta(c)$ and the shadow cost λc where λ is the multiplier of the budget constraint. The optimal rationing rule assigns the good to consumers where $\beta(c) - \lambda c > 0$. When β is concave, this means that low-cost consumers will be supplied by the public.

The private firm's profit-maximizing strategy is more complicated. The willingness to pay τ for consumer (w, c) is given by $U(w - \tau) + \beta(c) = U(w)$. Now τ depends on w and c, and is increasing in both arguments. At cost c, if the firm sets a price p, the demand is $\int_{p < \tau(w,c)} dF(w)$ and the profit is $[p - c] \int_{p < \tau(w,c)} dF(w)$. Profits may not be monotone, and may well be increasing in c; prices may not be monotone in cost either. The proofs do rely on monotonicity, and results may have to be modified. Nevertheless, our results in Sections 3 and 4 should continue to hold when the benefit function $\beta(c)$ is approximately a constant.

B.3. Correlation Between Wealth and Costs

The case of correlated wealth and costs shares the same issues as those in benefits dependent on costs. First, for the public supplier observing only wealth, we will have to consider the conditional distribution $G(c|w)$, and the conditional expected costs $\gamma(w) \equiv \int c \, dG(c|w)$. The budget constraint (4) becomes $\int_{\underline{w}}^{\overline{w}} \gamma(w)(1 - \theta(w)) f(w) \, dw \leq B$. The first-order derivative with respect to $\theta\bar{f}$ becomes $-1 + \lambda\gamma(w)$. Hence, if the conditional expected cost $\gamma(w)$ is increasing in w, the public supplier rations the rich consumers, and the converse is true if $\gamma(w)$ is decreasing in w. Second, the case of rationing by the public supplier observing both wealth and cost is unaffected by any correlation. The public supplier already observes both benefit and cost information, and applies the cost-effectiveness principle.

The private firm's profit function is now $[1 - F(w|c)][\tau(w) - c]$, where $F(w|c)$ is the conditional distribution of w given c. The profit function may not be decreasing in c and the profit-maximizing price may not be increasing in c. For example, suppose that w and c are negatively correlated. Then if the firm observes consumers with low costs, then these must be relatively wealthy consumers and the firm sets a high price. Again, prices may not be monotone in general, and our results may have to be modified. Nevertheless, prices decreasing with costs seem unrealistic. When prices are increasing in costs, our results in Sections 3 and 4 should be valid.

References

BARROS, P. P., and P. OLIVELLA (2005) Waiting lists and patient selection, *Journal of Economics & Management Strategy* **14**, 623–646.

BERNHEIM, B. D., and M. D. WHINSTON (1986) Common agency, *Econometrica* **54**, 923–942.

BESLEY, T., and S. COATE (1991) Public provision of private goods and the redistribution of income, *American Economic Review* **81**, 979–984.

BLACKORBY, C., and D. DONALDSON (1988) Cash versus kind, self-selection, and efficient transfers, *American Economic Review* **78**, 691–700.

GRASSI, S., and C. A. MA (2011) Optimal public rationing and price response, *Journal of Health Economics* **30**(6), 1197–1206.

HELLWIG, M. (2003) Public-good provision with many participants, *Review of Economic Studies* **70**, 589–614.

HOEL, M. (2007) What should (public) health insurance cover? *Journal of Health Economics* **26**, 251–262.

HOEL, M., and E. M. SÆTHER (2003) Public health care with waiting time: The role of supplementary private health care, *Journal of Health Economics* **22**, 599–616.

IVERSEN, T. (1997) The effect of a private sector on the waiting time in a national health service, *Journal of Health Economics* **16**, 381–396.

NORMAN, P. (2004) Efficient mechanisms for public goods with use exclusions, *Review of Economic Studies* **71**, 1163–1188.