Nonlinear Pricing for Average-Price-Biased Consumers: A Basis for Increasing Block Tariffs*

Phuong Ho†

August 19, 2020

Abstract

Although efficient nonlinear prices typically feature quantity discounts, policy-makers have been using increasing block tariffs to regulate utilities, relying on equity justification rather than efficiency. This paper re-examines regulatory nonlinear pricing in light of recent evidence that shows electricity consumers respond to changes in average prices rather than marginal prices. Results show pricing for biased consumers favors the distributional goals of the regulators. The optimal pricing features quantity premia rather than quantity discounts. Allocation favors low-type consumers due to efficiency at the bottom and distortion at the top. More people will be able to purchase a unit in the pricing scheme for average-price-biased consumers. Consequently, pricing in the biased consumer world may offer higher welfare, especially if most of the population is low type.

Keywords: nonlinear pricing, nonlinear Ramsey pricing, average-price bias, quantity premia

JEL Classification: D42, D82, D91

*A previous working version is "Nonlinear pricing, biased consumers, and regulatory policy" (May 13, 2017). I thank Stanley Reynolds, Derek Lemoine, Mauricio Varela, Gautam Gowrisankaran, Richard Green, and seminar participants at the University of Arizona, 2017 Cologne International Energy Summer School for comments and useful discussions.

†SNF – Centre for Applied Research at NHH, Helleveien 30, NO-5045 Bergen, Norway. Email: phuong.ho@snf.no.

1 Introduction

Nonlinear price schedules are observed for many goods and services. Firms offer quantity discounts on auto rentals, hotel room stays, and a wide variety of consumer goods. Cell phone service providers have complex nonlinear prices based on menus of three-part tariffs. These profit-maximizing nonlinear price schedules feature quantity discounts (decreasing marginal prices) for a wide variety of environments, with the highest consumer type paying a marginal price equal to the firm's marginal cost.

Regulated firms also employ nonlinear pricing to price electricity and water. However, they charge increasing block tariffs, in which consumers are charged higher marginal prices for higher consumed quantities. From an economic-efficiency standpoint, increasing-marginal-price schedule is a puzzle. Suppose the firm operates with economies of scale and the regulator wishes to set prices so as to maximize welfare subject to a break-even constraint for the firm. Then, second-best optimal prices require distortions away from the marginal cost, but these distortions will typically involve smaller mark-ups of the marginal price over the marginal cost for higher quantities, that is, decreasing marginal prices. This result follows the result in the context of non-regulated nonlinear pricing by Mirrlees (1971) and Maskin and Riley (1984).

Despite the debate on efficiency, increasing-marginal-price schedules are argued to promote distributional goals.[1] For example, when energy costs increased significantly in the 1970s and 1980s, many electric utilities raised marginal rates on high-consumption service tiers, while keeping marginal rates low on lower tiers to protect lower-income consumers. This increasing-marginal-price schedule points to the equity-efficiency trade-off for regulators, whereby distributional goals for pricing are achieved only by sacrificing efficiency. Borenstein (2012) provides evidence regarding the efficiency cost of increasing block tariffs used by California electric utilities, although he analyzes the distortion away from setting an alternative flat-rate schedule.

This paper offers a different perspective on the efficiency debate for increasing-marginal-price schedules for price regulation. This perspective is based on recent empirical evidence on biased consumer responses to complex nonlinear price schedules. Using the data on electricity consumption in California, Ito (2014) shows household customers respond to changes in average prices rather than marginal prices. In light of this evidence, I characterize the optimal nonlinear pricing when consumers make decisions using average prices.

I find that biased consumer responses to price changes can have dramatic effects on second-best optimal prices. For reasonable configurations of consumer preferences and distributions of consumer types, the second-best price schedule for average-price-biased consumers offers quantity premia, whereas the one for marginal-price rational consumers features quantity discounts. The markup when pricing biased consumers, in fact, always favors low-type consumers. Specifically, this markup offers

[1] For example, see Wichelns (2013), who also argues increasing block tariffs promote energy-conservation goals.

efficiency at the bottom and distortion at the top. Therefore, the marginal type pays the marginal tariff equal to the marginal cost of production of the unit. The highest type consumes less than the efficient level, paying the marginal tariff above the marginal cost. These results sharply contrast with the outcomes in pricing for rational consumers where efficiency is achieved at the top and distortion arises at the bottom.

The price schedule for average-price-biased consumers also favors the distributional goal of the regulators in which it expands the participating population. With zero fixed costs of production, the marginal type under the average-price-consumer pricing regime is the first-best level and thus lower than the one under the marginal-price-consumer regime. Therefore, more people are able to purchase units in the case of pricing-biased consumers.

With such favors toward low-type consumers and the expansion of the participating population, the nonlinear pricing for biased consumers offers welfare at least as equal as the pricing for marginal-price rational consumers. Welfare improvement happens if most of the population is low type (e.g., consumer heterogeneity is distributed according to Pareto distribution).

However, welfare loss may happen when the firm has fixed costs. The reason is that the firm is not able to pass any part of fixed costs to the tariff when pricing biased consumers (though it can pass a part of the fixed costs to the fixed fee of the tariff when pricing rational consumers). Instead, he recovers the fixed costs by shrinking the participating population. This raises the marginal type in the case of pricing for biased consumers to above the first-best level, and may even cause it to surpass the marginal type in the case of pricing rational consumers.

My model is built on the literature of nonlinear pricing in Mussa and Rosen (1978), Maskin and Riley (1984), and Stole (2007). I depart from the literature by assuming consumers respond to average prices instead of marginal prices. This setting is a natural one given recent evidence on this type of behavior. In markets with nonlinear price schedules such as utility pricing and taxation, De Bartolome (1995) shows tax payers use average tax rates. Borenstein (2009) and Saez (2010) show consumers are more likely respond to expected marginal prices than marginal prices. In a more comprehensive setting that uses a more general form of perceived price that covers three possible perceptions (marginal prices, expected marginal prices, and average prices), Ito (2014) provides strong empirical evidence of average-price responses in electricity consumption.[2]

A growing literature of behavioral economics in industrial organization also examines how firms react to and in some cases exploit consumers' suboptimal responses. These studies consider types of bias apart from average-price bias, such as loss aversion, present bias, overconfidence, or failure to choose the best price due to suboptimal search. Readers can refer to a comprehensive survey by DellaVigna (2009). For example, Courty and Hao (2000), Eliaz and Spiegler (2008), and Grubb (2009) study monopoly screening problems or price discrimination when consumers are overconfident. Heidhues and Kőszegi (2008) and Spiegler (2012) suggest loss aversion may create kinks in

[2]In a similar approach, Ito (2013) provides the evidence of average-price responses in water consumption.

demand curves, which can lead to price rigidities. Carbajal and Ely (2016) and Hahn et al. (2018) characterize price discrimination when a monopolist faces consumers with loss aversion and state-contingent reference points.

This paper is closely related to two early papers. Liebman and Zeckhauser (2004) model the optimal tax rate when people respond to average tax rates as an approximation of marginal prices due to the substantial cognitive cost of understanding the complex price schedules. They do not study the impact on the optimal tax shape. Furthermore, note the nonlinear pricing model in labor economics, although similar to other product and service markets, exhibits significant differences in optimization constraints and in how consumption preferences change in consumer heterogeneity.[3] Sobel (1984) considers naive "price-taking" consumers and explains average-price response results from consumers thinking they face a linear price schedule with a constant unit rate. He assumes constant average costs and that demand becomes more elastic as consumption increases. I focus on a more natural setting to see how the cost structures and preferences affect the firm's profits and consumer surplus. I emphasize that neither Liebman and Zeckhauser (2004) nor Sobel (1984) identify the important implication of consumers' average-price bias on nonlinear pricing. That is, nonlinear pricing in the world of biased consumers offers several features favoring distributional goals of regulators: convex tariff, efficiency at the bottom, distortion at the top, and first-best participation. Presenting these features and how they contribute to the change in social welfare in the regulated pricing scheme for biased consumers relative to the one in the regulated pricing scheme for rational consumers is the primary contribution of this paper.[4]

The rest of the paper is organized as follows. Section 2 discusses the reasons for average-price bias

[3]The monopolist in Liebman and Zeckhauser's model has a constant marginal cost of production, sells products to two-type consumer types, and is constrained to use a piecewise-linear price schedule originating from the origin. With this constraint, the monopolist always prefers selling to biased consumers to selling to marginal-price consumers. Farhi and Gabaix (2020) develop a general framework of optimal taxation with behavioral agents that encompasses the study of Liebman and Zeckhauser (2004) and other biases. For the implications of average-price bias on monopoly pricing, Liebman and Zeckhauser (2004) is more directly related to my paper. Also note the optimal shape of the income tax rate has been well studied in Mirrlees (1971), Diamond (1998), and Dahan and Strawczynski (2000). Most analysis of optimal taxation supports decreasing marginal tax rates at high levels of income (see summary in Dahan and Strawczynski (2000)). Diamond (1998) and the later comment by Dahan and Strawczynski (2000) demonstrate the utility form and the worker skill distribution (equivalent to the consumer type distribution in product and service markets) account for the optimal tax structures. Hence, the optimal tax rate may be increasing or U-shaped. This result, under the classic assumption that consumers respond to marginal prices, is not surprising, because the hazard rate of the distribution of the consumer type does affect the tariff shape, following a result in Maskin and Riley (1984). However, such satisfied type distribution is not popular, thus the obscurity of the result.

[4]Working independently, Martimort and Stole (forthcoming) take a similar approach and arrive at similar results. However, whereas they study the implications of bias for monopoly pricing, I consider the implications on regulatory pricing. I also cover how consumer heterogeneity affects social welfare. Including the effects of consumer heterogeneity is important, because it is a key point affecting the tariff shape under the classic assumption of marginal price response (see the results in Maskin and Riley, 1984). Furthermore, whereas Martimort and Stole emphasize the replacement of information rents for curvature rents, I add that the markup when pricing biased consumers is the inverse price elasticity of quantity demanded rather than the inverse price elasticity of population demanded in the case of rational consumers. Ultimately, the markup when pricing biased consumers is priced at the net willingness to pay of every type-specific consumer.

and the behavioral theory behind average-price response. Section 3 outlines the models of nonlinear pricing when consumers respond to marginal prices and when consumers respond to average prices. Section 4 shows the implications of the average-price response of consumers on the tariff structure and consumption allocation. Section 5 presents the comparison of welfare between the two consumer-response regimes. Section 6 concludes.

2 Rationing average-price bias of consumers' response

In the monopoly pricing problem, the firm designs the price schedule (the pairs of price rate and consumption quantity) that maximizes expected profits, given that consumers facing the price schedule will select the consumption that maximizes their individual utilities. This constraint is known as the incentive constraint whereby each type of consumer selects their own prescribed consumption and does not have an incentive to buy other consumptions (that are designed for other types). Formally, this incentive compatibility constraint is written as

$$q(\theta) \text{ solves } \max \mathcal{U}(q, \theta, P(q)) \text{ for each } \theta, \tag{1}$$

where a type-θ consumer demands $q(\theta)$ units to maximize her utility, taking account of the entire total price schedule $P(q)$.

Under the conventional utility function $\mathcal{U}(q, \theta, P(q)) = U(q, \theta) - P(q)$, the type-$\theta$ consumer compares the marginal utility with the marginal price and opts to consume units at which marginal utility equals marginal price, that is, $q(\theta)$ such that $U_q(q(\theta), \theta) = P_q(q(\theta))$ (subscripts denote derivatives).

In the world of average-price-response behavior, the type-θ consumer will demand $q(\theta)$ such that $U_q(q(\theta), \theta) = P(q(\theta))/q(\theta)$. The reason is that consumers either use average prices to approximate marginal prices or they misperceive the entire total price schedule.

Liebman and Zeckhauser (2004) propose that consumers treat average prices as an alternative approximation for marginal prices. They call it "ironing" behavior, and identify three possible reasons. First, complexity, for example, due to the nonlinear structure, makes determining marginal prices hard, and make knowing where one is on the price schedule costly. Compared with marginal price, much less information is required to calculate average price: Only the total payment and quantity are sufficient. Second, the connection between a consumer's choice and consumption is difficult to observe, especially in purchasing electricity and water. For example, how many kilowatts of electricity are used to cook a meal, and how much it costs to cook that meal is even more difficult to perceive. Third, a nonstationary environment is not conducive for consumer learning. Different seasons often lead consumers to stay at very different points on the price schedule (more heating is demanded in the winter). The pricing schedules are, however, not displayed on the monthly bills, whereas monthly payment aggregates hundreds of disparate single activities (turning on the light, running the refrigerator, using the heater, etc.).

The ironing manner is empirically supported by Ito (2014). He shows consumers do respond to prices, but they perceive average prices instead of marginal prices because they are inattentive to the price schedule. In his model, he recovers the perceived price $\tilde{p}(x)$ as a noise around the marginal price $p(x)$, that is, $\tilde{p}(x) = \int p(x-\epsilon)w(\epsilon)d\epsilon$. Using electricity-consumption data in California, the result shows the noise ϵ is uniformly distributed over $[0, x]$, which means the perceived price is the average price. Hence, consumers perceive the average price as a way to approximate the marginal price by weighting equally around it.

The second explanation for the average-price response is that consumers think they face a linear price schedule with a flat rate (marginal prices are also average prices). In a formal model by Sobel (1984), a consumer who demands $q(\theta)$ thinks she faces a linear price schedule with constant unit charge equal to $P(q(\theta))/q(\theta)$:

$$q(\theta) \text{ solves } \max \mathcal{U}(q, \theta, P(q(\theta))/q(\theta)). \tag{2}$$

Sobel (1984) shows the consumer's behavior converges dynamically toward the average-price response. The process suggests the reason is that the consumer knows only quantity and total payment of her purchase, not the entire price schedule. The consumer begins by making a purchase of q_0 and then learns about the value of $P(q_0)$. She makes the next demand under the assumption that the price is linear and the unit price is $P(q_0)/q_0$. In general, she makes her nth purchase assuming the average price of the $(n-1)$th purchase is the constant unit cost. This adjustment process converges to the problem (2).

A recent theory developed by Esponda and Pouzo (2016) explains average-price acting behavior as a result of a Berk-Nash equilibrium when agents make decisions in a misspecified environment. That is, instead of assuming people have a correctly specified view of their environment, each player is characterized by a subjective model that describes the set of feasible beliefs over payoff-relevant consequences as a function of actions. In a Berk-Nash equilibrium, each player follows a strategy that is optimal given her best-fit belief. The notion of best-fit belief is formalized in terms of minimizing the Kullback-Leibler divergence, which is endogenous and depends on the equilibrium strategy profile.

Applying their model to the average-price-response context, the true environment is the set of all possible unit costs $\mathbb{Y} = \{y \in \mathbb{R} : y = p(q)/q, q \in \mathbb{R}\}$. The consumer incorrectly believes she faces a (possibly random) linear price $y \in \mathbb{Y}$ that does not depend on her choice. Her misspecified subjective model is the set of all probability distributions over $\mathbb{Y}$. That is, $\Theta = \Delta(\mathbb{Y})$ where $\theta = (\theta_1, \ldots, \theta_k) \in \Theta$ and θ_j denotes the probability that the linear price is $P(q_j)/q_j$. The consumer's strategy is denoted by $\sigma = (\sigma_1, \ldots, \sigma_k) \in \Delta(\mathbb{Q})$, where σ_j is the probability that the consumer chooses quantity q_j. Esponda and Pouzo (2016) show that for a strategy σ, the weighted Kullback-Leibler divergence function is

$K(\sigma,\theta)=\sum_{j=1}^{k}\sigma_j \ln\frac{1}{\theta_j}$ and the unique minimizer is $\theta(\sigma)=(\sigma_1,\ldots,\sigma_k)$. Given such best-fit belief, the strategy σ maximizes the consumer's surplus $\{U(q)-\left(\sum_{j=1}^{k}(P(q_j)/q_j)\sigma_j\right)q\}$. Thus, σ is a Berk-Nash equilibrium. This strategy σ includes a special case of σ that defines a pure strategy q_j and the belief of the linear price with unit cost $P(q_j)/q_j$ with probability of one. In other words, average-price response is a result of a pure-strategy Berk-Nash equilibrium.

In this paper, I do not model the rationality of average-price-response behavior. Instead, I consider the design for the optimal price schedule of a monopoly and of a regulator, taking the average-price response of consumers exogenously. Note that consumers are biased in a way that they perceive average prices instead of marginal prices, but they are partly rational in a manner that they compare marginal utility with average prices rather than average utility with average prices. The perception of marginal utility is not unusual, because consumers are not completely irrational. They misunderstand only the price schedule rather than their true preferences.

3 Model

3.1 Pricing for marginal-price rational consumers (a review)

To begin with, I review the standard model in which consumers make choices based on marginal prices. This benchmark model is introduced in textbooks, following Mussa and Rosen (1978), Maskin and Riley (1984), and Stole (2007). In this model, a type-θ consumer gets the payoff

$$U(q,\theta)-P(q), \tag{3}$$

where $q\in\mathcal{Q}\subseteq\mathbb{R}_+$ is the quantity consumed; $\theta\in\Theta=[\theta_0,\theta_1]$ is a one-dimensional type distributed according to differentiable distribution $F(\theta)$, density $f(\theta)$; $P(q)$ is the total payment made to the firm for q units. The outside option of not consuming is assumed to be zero, $U(0,\theta)=0$. Denote subscripts as partial derivatives. Standard assumptions for consumer preferences specify U is thrice continuously differentiable, strictly concave, nondecreasing in q and θ ($U_q\geq 0, U_\theta\geq 0$), and $U_{q\theta}>0$. The last assumption is the familiar single-crossing property that implies higher-type consumers have higher marginal values of consumption.

Assuming the firm has convex cost $C(q)$, the profit-maximizing monopolist offers a nonlinear price schedule P: $\mathcal{Q}\to\mathbb{R}_+$ that maximizes

$$\max_{P(\cdot)}\int_{\theta_0}^{\theta_1}\{P(Q^*(\theta))-C(Q^*(\theta))\}\,dF(\theta) \tag{4}$$

$$\text{subject to }\; Q^*(\theta)=\arg\max_{q\geq 0}\; U(q,\theta)-P(q) \qquad \text{(incentive compatibility)} \tag{5}$$

$$U(Q^*(\theta))-P(Q^*(\theta))\geq 0 \qquad \text{(individual rationality).} \tag{6}$$

To distinguish the price schedule designed for rational consumers from the one for average-price biased consumers, I denote X^* and $\overline{X}$ as value X for the scenarios in which consumers respond to marginal prices and average prices, respectively.

The incentive compatibility constraint ensures that people are motivated to behave in a manner consistent with the optimal solution. Thus, a consumer with a higher type must consume his true preferred quantity rather than pretending to be a low type. For this reason, this constraint is also called the "truth telling condition." The other individual rationality constraint means people want to consume; that is, they are at least as well off by consuming as they would by not purchasing.

To solve for the optimal tariff, we rewrite the monopolist's profits as the difference between social welfare and consumer surplus:

$$\int_{\theta_0}^{\theta_1} \left\{ \underbrace{(U(Q^*(\theta), \theta) - C(Q^*(\theta)))}_{\text{social welfare } \equiv W^*(\theta)} - \underbrace{(U(Q^*(\theta), \theta) - P(Q^*(\theta)))}_{\text{consumer surplus } \equiv CS^*(\theta)} \right\} dF(\theta) \tag{7}$$

and notice that using integration by parts, the ex-ante consumer surplus equals

$$ECS^* \equiv \int_{\theta_0}^{\theta_1} CS^*(\theta) dF(\theta) \tag{8}$$

$$= \int_{\theta_0}^{\theta_1} (1 - F(\theta)) \cdot \frac{dCS^*(\theta)}{d\theta} \cdot d\theta + CS^*(\theta_0). \tag{9}$$

Because CS^* is the net indirect utility, that is, $CS^*(\theta) \equiv \max_{q \geq 0} \ U(q, \theta) - P(q)$, the additional surplus of a higher type is

$$\frac{dCS^*(\theta)}{d\theta} = U_\theta(Q^*(\theta), \theta). \quad \text{(envelope theorem's result)} \tag{10}$$

Therefore, the monopolist's expected profit is a tradeoff between efficiency and information rent:

$$\int_{\theta_0}^{\theta_1} \underbrace{(U(Q^*(\theta), \theta) - C(Q^*(\theta)))}_{\text{social welfare}} dF(\theta) - \int_{\theta_0}^{\theta_1} \underbrace{\frac{1 - F(\theta)}{f(\theta)} \cdot U_\theta(Q^*(\theta), \theta)}_{\text{information rent}} dF(\theta) - CS^*(\theta_0). \tag{11}$$

Assuming the allocation of $Q^*(\theta)$ is nondecreasing, the monopolist's profit-maximization problem is equivalent to finding the optimal allocation that maximizes the virtual profit:

$$Q^*(\theta) = \arg\max_{q \in \mathcal{Q}} \int_{\theta_0}^{\theta_1} \underbrace{U(Q^*(\theta), \theta) - C(Q^*(\theta)) - \frac{1 - F(\theta)}{f(\theta)} \cdot U_\theta(Q^*(\theta), \theta)}_{\text{virtual profit}} dF(\theta). \tag{12}$$

So, the optimal allocation $Q^*(\theta)$ must satisfy:

$$U_q(Q^*(\theta),\theta) - C'(Q^*(\theta)) = \frac{1-F(\theta)}{f(\theta)} \cdot U_{q\theta}(Q^*(\theta),\theta). \tag{13}$$

Then, the corresponding optimal tariff, $P^*(q)$, is uniquely defined from the marginal condition $P'(q) = U_q(q, \vartheta^*(q))$ and the participation constraint $P(Q^*(\theta^*)) = U(Q^*(\theta^*))$, where $\vartheta^*(q)$ is the inverse function of the allocation $Q^*(\theta) = q$ and θ^* is the marginal type at which virtual profit is zero.

3.2 Pricing for average-price-biased consumers

In the pricing regime for marginal-price rational consumers, the truth-telling condition is motivated by the fact that every consumer's true preferred consumption maximizes her own surplus. This condition is accomplished by choosing a consumption such that marginal utility equals marginal price. However, when consumers respond to average prices (either due to misperception of nonlinear prices as linear prices or imperfectly learning the complex structure of nonlinear prices; see section 2), they will choose a consumption such that marginal utility equals average prices. That is, the incentive compatibility constraint is implemented by $\overline{Q}(\theta) : U_q(q,\theta) = P(q)/q$. The consumer surplus hence equals

$$\overline{CS}(\theta) \equiv U(\overline{Q}(\theta),\theta) - P(\overline{Q}(\theta)) \tag{14}$$

$$= U(\overline{Q}(\theta),\theta) - \overline{Q}(\theta) \cdot U_q(\overline{Q}(\theta),\theta) \tag{15}$$

$$= \overline{Q}(\theta) \left[\frac{U(\overline{Q}(\theta),\theta)}{\overline{Q}(\theta)} - U_q(\overline{Q}(\theta),\theta) \right]. \tag{16}$$

Because U is strictly concave, average utility is greater than marginal utility. Consumer surplus is hence positive, and the participation constraint is trivially satisfied. To reflect that this surplus is positive owing to diminishing marginal utility, Martimort and Stole (forthcoming) refer to the surplus as the consumer's *curvature rent*. They also note the monopolist's profit is a tradeoff between efficiency and the curvature rent:

$$E\Pi = \int_{\theta_0}^{\theta_1} \underbrace{(U(\overline{Q}(\theta),\theta) - C(\overline{Q}(\theta)))}_{\text{social welfare}} dF(\theta) - \int_{\theta_0}^{\theta_1} \underbrace{U(\overline{Q}(\theta),\theta) - \overline{Q}(\theta) \cdot U_q(\overline{Q}(\theta),\theta)}_{\text{curvature rent}} dF(\theta). \tag{17}$$

Here, I highlight how the curvature rent and the information rent differ. To see their relation, let us go back one step and recall the definition of the ex-ante consumer surplus:

$$\begin{aligned} ECS &\equiv \int_{\theta_0}^{\theta_1} CS(\theta) dF(\theta) && (18) \\ &= \int_{\theta_0}^{\theta_1} (1 - F(\theta)) \cdot \frac{dCS(\theta)}{d\theta} \cdot d\theta + CS(\theta_0). \quad \text{(integration by parts)} && (19) \end{aligned}$$

Given $CS(\theta) \equiv U(Q(\theta), \theta) - P(Q(\theta))$, the additional surplus for a higher type is

$$\frac{dCS(\theta)}{d\theta} = U_\theta(q, \theta) + Q'(\theta) \cdot \left[U_q(Q(\theta), \theta) - P'(Q(\theta))\right]. \tag{20}$$

When consumers respond to marginal prices, the additional surplus for a higher type is just the additional utility of the higher type: $\frac{dCS^*(\theta)}{d\theta} = U_\theta(Q^*(\theta), \theta)$.

However, when consumers respond to average prices,

$$\frac{dCS(\theta)}{d\theta} = U_\theta(q, \theta) + Q'(\theta) \cdot \underbrace{\left[U_q(Q(\theta), \theta) - P'(Q(\theta))\right]}_{\neq 0} \tag{21}$$

$$\frac{d\overline{CS}(\theta)}{d\theta} = U_\theta(\overline{Q}(\theta), \theta) + \overline{Q}'(\theta) \cdot \left[AP(\overline{Q}(\theta)) - P'(\overline{Q}(\theta))\right] \quad \text{because } U_q = P/q \tag{22}$$

$$\text{or} \quad \frac{d\overline{CS}(\theta)}{d\theta} = U_\theta(\overline{Q}(\theta), \theta) - \underbrace{\overline{Q}(\theta) \cdot AP'(\overline{Q}(\theta)) \cdot \overline{Q}'(\theta)}_{\text{average-price bias loss}} \quad \text{because } P' = AP + qAP' = U_q + qAP', \tag{23}$$

where AP is average price.

Therefore, the information rent is offset by the loss due to average-price-biased behavior, leaving the remainder as the curvature rent. This sign of the average-price bias loss also reveals that the monopolist likely sets an increasing-price-rate tariff (to induce a positive average-price bias loss in equation (23)) to exploit the curvature rent. We discuss the tariff shape in more detail in section (4).

Recall our goal is to find the optimal nonlinear price schedule that maximizes the firm's profit subject to consumer response behavior. Because $P(q) = qU_q(q, \theta)$, the firm's problem is equivalent to finding the optimal allocation $\overline{Q}(\cdot)$ such that (assuming the allocation is increasing in θ, $\overline{Q}' > 0$)

$$\overline{Q}(\theta) = \arg\max_{q \in \mathcal{Q}} \int_{\theta_0}^{\theta_1} (qU_q(q, \theta) - C(q)) dF(\theta) \tag{24}$$

$$\Leftrightarrow \overline{Q}(\theta) = \arg\max_{q \in \mathcal{Q}} \; qU_q(q, \theta) - C(q). \tag{25}$$

The corresponding optimal tariff, $\overline{P}(q)$, is uniquely defined from the average-price condition $\frac{P(q)}{q} =$

$U_q(q, \overline{\vartheta}(q))$, where $\overline{\vartheta}(q)$ is the inverse function of $\overline{Q}(\theta) = q$. Marginal type $\overline{\theta}$ is determined by the break-even profit $P(\overline{Q}(\overline{\theta})) - C(\overline{Q}(\overline{\theta})) = 0$.

Proposition 1. *The optimal allocation when pricing average-price-biased consumers, $\overline{Q}(\theta)$, satisfies*

$$U_q(\overline{Q}(\theta), \theta) - C'(\overline{Q}(\theta)) = -\overline{Q}(\theta) \cdot U_{qq}(\overline{Q}(\theta), \theta), \tag{26}$$

and the optimal price schedule is

$$\overline{P}(q) = q \cdot U_q(q, \overline{\vartheta}(q)), \tag{27}$$

where $\overline{\vartheta}(q)$ are inverse functions of $\overline{Q}(\theta)$.

The proposition implies the difference in mark-up formulas when pricing for marginal-price-responding consumers versus for average-price-biased consumers:

$$\frac{MP^* - MC^*}{MP^*} = \frac{1 - F(\theta)}{f(\theta)} \frac{U_{q\theta}(Q^*, \theta)}{U_q(Q^*, \theta)} \tag{28}$$

$$\text{versus} \quad \frac{\overline{AP} - \overline{MC}}{\overline{AP}} = \frac{-\overline{Q} U_{qq}(\overline{Q}, \theta)}{U_q(\overline{Q}, \theta)}, \tag{29}$$

where AP, MP, and MC denote average price, marginal price, and marginal cost, respectively.

Two notable observations about these markups should be made. First, the markup in the marginal-price regime is marginal-price markup while the one in the average-price regime is average-price markup. This change is evident because the markup should be designed corresponding to the price rate to which consumers respond. Second, the markup in the average-price regime does *not* depend on the distribution of consumer types. Of course, this independence results from the fact that the monopolist in the marginal-price-responding case ultimately maximizes expected virtual profit, whereas the profit-maximization problem in the average-price-biased-consumer case turns out to be a point-wise maximization (25). However, this point-wise maximization does not mean the monopolist facing average-price-biased consumers can perfectly discriminate between consumer types. The point-wise maximization happens because the monopolist must exploit average-price bias universally, over *all* types of consumers. By contrast, in the marginal-price-response case, only high-type consumers receive information rents; thus, the monopolist facing marginal-price-responding consumers only needs to extract the surplus of a *portion* of the population.

As a further reflection of the notion of type independence versus a portion of population, the right hand sides of the above markup formulas are both inverse elasticities but have distinct meanings. To begin with, consider the purchase decision of a consumer, given type θ, when he responds to average prices. He is indifferent between buying and not buying the qth unit by comparing the marginal utility with the average price p^a of that unit:

$$U_q(q,\theta) = p^a \tag{30}$$

$$\Rightarrow U_{qq} dq = dp^a \tag{31}$$

$$\Rightarrow \frac{-qU_{qq}}{U_q} = -\frac{dp^a}{dq} \cdot \frac{q}{p^a} \equiv \frac{1}{\eta}, \tag{32}$$

which means $\frac{\overline{AP}-\overline{MC}}{\overline{AP}} = \frac{1}{\eta}$.

By contrast, given a quantity q, consider the demand for the qth unit of consumption in the marginal-price regime. This unit has the price p^m, which is the marginal price because consumers are responding to marginal price. In other words, we are considering the marginal-price demand in this case. The proportion of the population who is willing to buy this unit is

$$D(p^m) \equiv 1 - F(\theta^*(p^m)) \tag{33}$$

$$\Rightarrow \frac{dD(p^m)}{dp^m} = f(\theta^*(p^m))\frac{d\theta}{dp^m}, \tag{34}$$

where $\theta^*(p^m)$ denotes the type of consumer who is indifferent between buying and not buying the qth unit at marginal price p^m. This consumer, as assumed, makes a decision on marginal prices:

$$U_q(q, \theta(p^m)) = p^m \tag{35}$$

$$\Rightarrow U_{q\theta} d\theta = dp^m \tag{36}$$

$$\Rightarrow \frac{d\theta}{dp^m} = \frac{1}{U_{q\theta}} \tag{37}$$

$$\text{Hence:} \quad \frac{dD(p^m)}{dp^m} \cdot \frac{p^m}{D(p^m)} = \frac{f(\theta^*(p^m))}{1 - F(\theta^*(p^m))} \cdot \frac{p^m}{U_{q\theta}} \equiv \frac{1}{\epsilon}. \tag{38}$$

Note $p^m = U_q$, so $\frac{f(\theta^*(p^m))}{1-F(\theta^*(p^m))} \cdot \frac{U_q}{U_{q\theta}}$ is the inverse elasticity, or $\frac{MP^* - MC^*}{MP^*} = \frac{1}{\epsilon}$.

Consequently, whereas the markup ratio under pricing for marginal-price rational consumers is the inverse elasticity of population demanded, the markup ratio when pricing biased consumers is the inverse elasticity of quantity demanded. Therefore, the tariff in the case of average-price consumers is independent of consumer heterogeneity. In section 4, we see that the markup when pricing biased consumers is priced at the *net willingness to pay* of every type-specific consumer.

3.3 Regulated tariffs

A monopolist is often regulated by the price at which total welfare is maximized subject to the monopolist's profit at a fixed rate, often a break-even rate to just cover enough fixed costs c_0. This pricing method is also called Ramsey pricing in multiproduct contexts. I now characterize the optimal nonlinear pricing of a regulator for cases of rational consumers and average-price-biased consumers, respectively. The results are that the regulated tariffs feature general forms, covering special cases of monopoly pricing and first-best pricing.

First, consider the world of marginal-price rational consumers. The regulator's problem is to maximize social welfare subject to a constraint on the firm's profit, in addition to incentive compatibility and individual rationality constraints:

$$\max_{P(\cdot)} \int_{\theta_0}^{\theta_1} U(Q^*(\theta),\theta) - C(Q^*(\theta))dF(\theta) \tag{39}$$

$$\text{subject to} \quad \int_{\theta_0}^{\theta_1} P(Q^*(\theta)) - C(Q^*(\theta))dF(\theta) = c_0 \;\; \text{(Profit constraint)} \tag{40}$$

$$Q^*(\theta) = \arg\max_{q\geq 0} U(q,\theta) - P(q) \;\; \text{(incentive compatibility)} \tag{41}$$

$$U(Q^*(\theta)) - P(Q^*(\theta)) \geq 0 \;\; \text{(individual rationality).} \tag{42}$$

To solve this problem, we rewrite the profit as the difference between social welfare and consumer surplus in the profit constraint (40). Then we use the Lagrangian method to solve the optimization with profit constraint. Denoting λ^* as the multiplier associated with the profit constraint, the first-order condition of the Lagrangian function is

$$U_q - C_q = \left(\frac{\lambda^*}{1+\lambda^*}\right) \cdot \left(\frac{1-F}{f} \cdot U_{q\theta}\right). \tag{43}$$

The Lagrangian multiplier λ^* is the marginal increase in welfare associated with a decrease in firm profit. The first-order condition hence implies

$$\frac{MP^* - MC^*}{MP^*} = \frac{r^*}{\epsilon^*}, \tag{44}$$

where $r^* \equiv \frac{\lambda^*}{1+\lambda^*} \in (0,1)$. This multiplier is the *regulating degree*, ruling the first-best case ($r \to 0$), regulated case ($0 < r < 1$), and monopoly case ($r \to 1$).

Similarly, when consumers respond to average prices, the regulator also maximizes social welfare subject to the firm's profit constraint, the average-price-response condition, and the individual rationality (trivially satisfied owing to concave utility). The result is that the optimal mark-up satisfies

$$\frac{\overline{AP} - \overline{MC}}{\overline{AP}} = \frac{\overline{r}}{\overline{\eta}}, \tag{45}$$

where $\overline{r}$ is the regulating degree for the average-price-bias case, $\overline{r} \equiv \frac{\overline{\lambda}}{1+\overline{\lambda}} \in (0,1)$.

In summary, the two response behaviors of consumers result in remarkable difference in markups in the firm's and the regulator's pricing. In the marginal-price-response regime, the markup is gained from the information rent of the high types. In the average-price-response regime, the markup is a result of exploiting the bias of all consumer types. Pricing for biased consumers of course has a different tariff shape from the schedule for rational consumers. The tariff for biased consumers is likely convex in order for the firm to exploit the curvature rent by extracting a positive average-price-bias loss (see equation (23)). In the next section, we see that the convex tariff shape in pricing biased consumers happens because this pricing features efficiency at the bottom and distortion at the top, and the markup prices the net willingness to pay of every consumer.

4 Tariff shapes and consumption distribution

In the literature on nonlinear pricing as well as second-degree price discrimination, efficient tariffs are known to offer quantity discounts. Consumers are charged lower marginal prices for higher consumed quantities. This result relies on certain type distributions of consumer types. A well-known illustration is from Maskin and Riley (1984), who relate the tariff shapes to the hazard rate of the consumer types:

Remark. Assume the utility form $U(q,\theta) = \theta V(q)$. When the consumers respond to marginal price, the optimal pricing schedule offers quantity discounts if the hazard rate (or inverse Mill's ratio) of the type distribution is increasing.

Examples of an increasing hazard rate include regular distributions such as uniform, Gompertz, Weibull, Rayleigh, and so on.[5] This result implies regulators sacrifice efficiency when using increasing block tariffs to price utilies (electricity, water). These tariffs are controversial although the regulators support increasing block tariffs on grounds of distributional goals, by offering low marginal rates on lower consumption tiers to protect lower income consumers.

Such controversy no longer remains if we live in the world of average-price-biased consumers. As consumers respond to average prices, the optimal nonlinear pricing scheme will feature quantity premia, allowing increasing block tariffs to achieve both efficient and distributional goals regardless of the distribution of consumer types.

Corollary 1. *The optimal pricing schedule for average-price-biased consumers does not depend on the distribution of consumer types. Moreover, it offers quantity premia as long as the relative curvature* $\frac{-qU_{qq}}{U_q}$ *is increasing.*

To understand the intuition of the tariff shapes, looking at the consumption allocation between the two price-response regimes is useful. The mark-up formulas (28) and (29) in section (3.2) imply a remarkable difference in pricing efficiency between the two regimes. Specifically, the optimal nonlinear

[5] Distribution with decreasing hazard rate such as Pareto distribution would imply the decreasing marginal rate pricing schedule.

price schedule for marginal-price rational consumers offers efficiency at the top; that is, the highest consumer type pays a marginal tariff equal to the marginal cost of producing the last unit, and offers positive markups for the other consumer types. This attempt, to balance between efficiency and information rents, leads to a decreasing marginal-price scheme, assuming constant marginal cost. On the other hand, pricing for average-price-biased consumers offers *efficiency at the bottom* and *distortion at the top*: Marginal type is priced at marginal cost and other types are priced with positive markups. The efficiency at the bottom and distortion at the top results in an increasing average price scheme, assuming constant marginal cost.

A very clear and interesting way to understand the differences in markups and tariff shapes between the two regimes is to consider the optimal price schedules using a two-type example. Using the discrete two-type case, we see that when consumers respond to average prices, the monopolist prices the markup at the *net willingness to pay* of every consumer; that is, the term $-qU_{qq}$ in the markup formula is the willingness to pay net of the average price.

As an illustration, consider that a monopolist offers two bundles: $(q_1, P(q_1))$, which is directed at type-1 (θ_1) consumers (in proportion ρ of the population), and $(q_1, P(q_1))$, which is directed at type-2 ($\theta_2 > \theta_1$) consumers (in proportion $1-\rho$). The monopolist's profit is

$$\Pi = \rho(P(q_1) - C(q_1)) + (1-\rho)(P(q_2) - C(q_2)) \tag{46}$$

subject to the individual rationality constraint; that is, the net surplus of the low type consumers must be positive:

$$U(q_1, \theta_1) - P(q_1) \geq 0. \tag{47}$$

It is also subject to the incentive compatibility constraints; that is, none of the types have an incentive to consume the other type's bundle. In particular, the high-type consumers should not want to consume the low-type consumers' bundle:

$$U(q_2, \theta_2) - P(q_2) \geq U(q_1, \theta_2) - P(q_1). \tag{48}$$

We skip the other incentive compatibility constraint for the low type because it is automatically implied by the two above constraints. The monopolist's problem is equivalent to choosing (q_1, q_2) to maximize the profits in the expression 46 subject to the constraints 47 and 48. Because the monopolist benefits from high prices, constraint 47 implies $P(q_1) = U(q_1, \theta_1)$. Constraint 48 then implies $P(q_2) = U(q_2, \theta_2) - (U(q_1, \theta_2) - U(q_1, \theta_1))$. These two expressions mean the monopolist chooses the price $P(q_1)$ so as to appropriate the type-1 consumers' surplus entirely. Meanwhile, the choice of $P(q_2)$ leaves some net surplus to the type-2 consumers because they can always buy the bundle $(q_1, P(q_1))$ and have the net surplus $U(q_1, \theta_2) - P(q_1) = U(q_1, \theta_2) - U(q_1, \theta_1)$, which is the information rent.

To formally get the markup, substitute the constraints into the objective function and rewrite the monopolist's profit maximization:

$$\max_{q_1,q_2} \rho[U(q_1,\theta_1) - C(q_1)] + (1-\rho)[U(q_2,\theta_2) - C(q_2) - (U(q_1,\theta_2) - U(q_1,\theta_1))]. \tag{49}$$

Denoting the marginal utility of type 1 by $MU(q_1,\theta_1)$, that is, $\partial U(q_1,\theta_1)/\partial q_1$, marginal price by MP and marginal cost by MC, the two first-order conditions (FOCs) are

$$\rho[MU(q_1,\theta_1) - MC(q_1)] = (1-\rho)[MU(q_1,\theta_2) - MU(q_1,\theta_1)] \tag{50}$$

$$MU_2(q_2) = MC(q_2). \tag{51}$$

Recall that the full rationality of consumers refers to $MU(q_1,\theta_1) = MP(q_1)$ and $MU(q_2,\theta_2) = MP(q_2)$. Hence, the second FOC means pricing for marginal-price rational consumers is efficient at the top; that is, the high-type (type 2) consumer pays the marginal tariff equal to the marginal cost of the unit. Meanwhile, the first FOC implies the monopolist has to balance the extraction of information rents with providing consumers with the right incentives not to imitate other consumer types:

$$\underbrace{\rho[MP(q_1) - MC(q_1)]}_{\text{expected profits by selling } q_1 \text{ to low type}} = \underbrace{(1-\rho)[MU(q_1,\theta_2) - MP(q_1)]}_{\text{opportunity cost of giving up the high type's higher willingness to pay for } q_1}. \tag{52}$$

This first FOC also means the monopolist has to equalize the gains of selling q_1 units at the low type's willingness to pay and at the high type's willingness to pay:

$$\underbrace{\overbrace{\rho[MU(q_1,\theta_1) - MC(q_1)]}^{\text{profits by selling } q_1 \text{ to } \theta_1} + \overbrace{(1-\rho)[MU(q_1,\theta_1) - MC(q_1)]}^{\theta_2 \text{ only pays this amount for the unit } q_1}}_{\text{expected profits if designing the price } MP(q_1)=MU(q_1,\theta_1)} =$$

$$= \underbrace{\overbrace{0}^{\theta_1 \text{ can't afford}} + \overbrace{(1-\rho)[MU(q_1,\theta_2) - MC(q_1)]}^{\text{profits by selling only to } \theta_2}}_{\text{expected profits if designing the price } MP(q_1)=MU(q_1,\theta_2)} \tag{53}$$

The monopolist's cost of giving up the high type's willingness to pay for q_1 becomes information rent for the high-type consumer.

Now, let us consider the monopolist's problem when consumers respond to average prices. In this scenario, the incentive compatibility constraints are replaced by

$$MU(q_1,\theta_1) = \underbrace{U(q_1,\theta_1) - 0}_{\text{additional utility from zero consumption to consuming } q_1} \geq AP(q_1) = \frac{P(q_1)}{q_1} \tag{54}$$

$$MU(q_2,\theta_2) = \underbrace{U(q_2,\theta_2) - U(q_1,\theta_2)}_{\text{additional utility from consuming } q_1 \text{ to consuming } q_2} \geq AP(q_2) = \frac{P(q_2)}{q_2}. \tag{55}$$

This replacement happens because consumers compare the marginal utility with the average price to decide whether they consumer the amount of good. Consumers consume their true preferences because they misperceive that average price is a flat rate and that the price rate does not depend on the consumption units. That is, consumers misperceive the price schedule as $P(q) = y \cdot q$, where y is a flat average rate. Given the price rate does not depend on consumption units, consumers do not have an incentive to imitate the other consumption.

Using these two contraints, rewrite the monopolist's profit-maximization problem:

$$\max_{q_1,q_2} \rho[q_1 MU(q_1,\theta_1) - C(q_1)] + (1-\rho)[q_2 MU(q_2,\theta_2) - C(q_2)]. \tag{56}$$

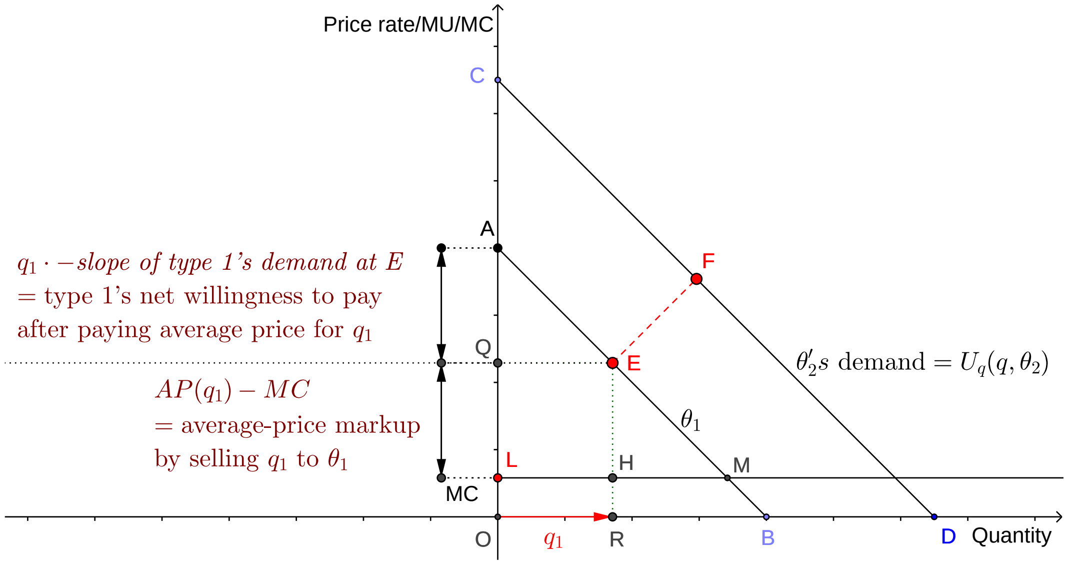
The first-order conditions are

$$\begin{aligned} &\rho[MU(q_1,\theta_1) + q_1 MU_q(q_1,\theta_1) - MC(q_1)] = 0 \\ \Leftrightarrow \quad &\underbrace{AP(q_1) - MC(q_1)}_{\text{monopolist's mark up of bundle } (q_1, P(q_1))} = \underbrace{-q_1 MU_q(q_1,\theta_1)}_{\text{type 1's net willingness to pay for } q_1} \end{aligned} \tag{57}$$

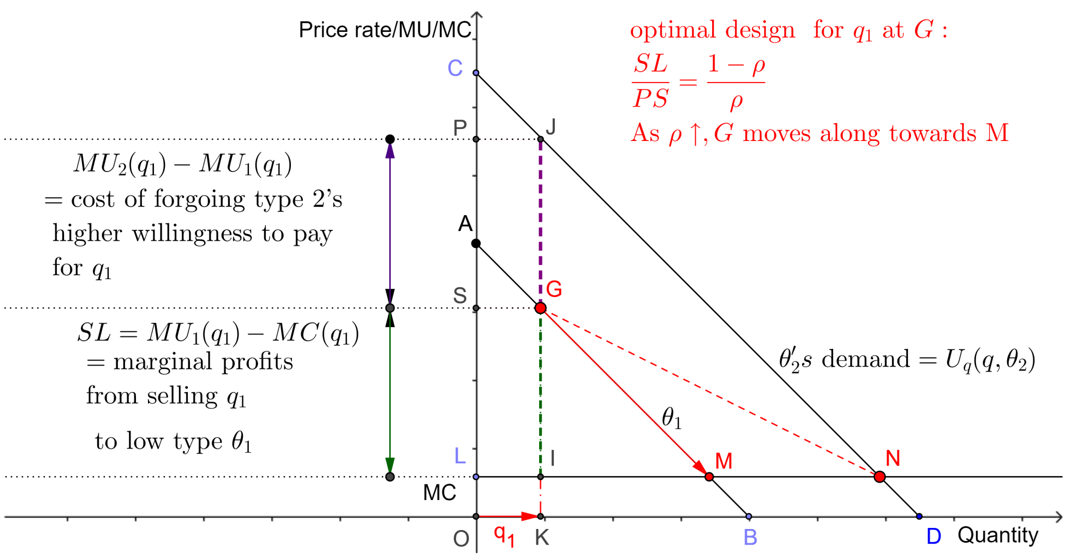
$$\begin{aligned} \text{and, } &\rho[MU(q_2,\theta_2) + q_2 MU_q(q_2,\theta_2) - MC(q_2)] = 0 \\ \Leftrightarrow \quad &\underbrace{AP(q_2) - MC(q_2)}_{\text{monopolist's mark up of bundle } (q_2, P(q_2))} = \underbrace{-q_2 MU_q(q_2,\theta_2)}_{\text{type 2's net willingness to pay for } q_2} . \end{aligned} \tag{58}$$

Therefore, the markup offers efficiency at the bottom; that is, only the marginal type (with consumption of zero units) pays the marginal costs of production. All other types of consumers, as long as they are purchasing a positive quantity, pay positive markups equal to their *net willingness to pay*.

To see why the right-hand-side term is the net willingness to pay for the units, let us consider Figure 1. This figure illustrates the optimal pricing schedules between the two response regimes in the discrete two-type consumer case with linear demands. Panel A shows the optimal pricing scheme for average-price-biased consumers. The marginal type that consumes zero units is priced at marginal cost, point L. The other consumer types are priced at positive markups, at points E and F for types θ_1 and θ_2, respectively. These markups are set to balance with the net willingness to pay of the corresponding consumer type. For example, type θ_1 is priced at point E. She pays the markup $AP(q_1) - MC$, which is the segment QL in the figure. She is willing to pay this markup because



(A) When consumers respond to average prices, optimal price rates are E and F.



(B) When consumers respond to marginal prices, optimal price rates are G and N.

FIGURE 1: Monopolist's optimal price schedule under average-price-response scheme vs. under marginal-price-response scheme in the case of two types of consumers, θ_1 and θ_2.

Note: The horizontal axis is for quantity q, the vertical axis is for the responding price rate p (average price or marginal price), or marginal utility MU, or marginal cost MC. Under average-price-response regime, the monopolist's optimal (average) price rate includes E and F. The optimal rate E is designed such that the mark up price equals the net willingness to pay after paying average price for the quantity q^E. Efficiency happens at the bottom, that is, at the marginal type at point L. Under marginal-price-response regime, the monopolist's optimal (marginal) price rate includes G and N. Efficiency happens at the top, that is, at the highest type at point N. If the low type and the high type occupy the population equally, the optimal choice G is designed such that the markup from selling q_1 to type θ_1 equals the opportunity cost of forgoing the marginal willingness to pay of the high type θ_2. When the population is more likely to be the low type (ρ increases), letting the low type consumer purchase more is more profitable; that is, optimal G moves along towards M.

it equals $-q_1 MU_q(q_1, \theta_1)$, or $q_1 \cdot$ *the slope of the demand of type 1*, which is her *willingness to pay net off the average price rate*, represented by the segment AQ. Designing the price schedules that starts at zero markup for the marginal consumer type and positive markups that equal net willingness to pay leads to an increasing average-price-rate schedule for the monopolist. Because consumers' misperception about the pricing structures gives them no incentive to imitate nearby consumer types, the monopolist does not have to balance the extraction of the information rent, resulting in a price schedule that does not depend on the distribution of consumer types.

Panel B shows the optimal pricing scheme for marginal-price rational consumers. As we have known, the monopolist finds it optimal to sell as much as possible to the high type θ_2, that is, efficiency at the top, point N. Meanwhile, the design for the optimal q_1 requires the monopolist consider the trade-off between selling to type 1 and selling to type 2. The monopolist can choose to serve the low type and get the marginal profits $MU_1(q_1) - MC(q_1)$. Such choice leaves out the high type θ_2's additional willingness to pay for unit q_1 by the amount $MU_2(q_1) - MU_1(q_1)$. Hence, the monopolist opts to choose q_1 such that the expected profits from selling q_1 to type 1 equal the opportunity costs of giving up the marginal valuation of the high type. This cost becomes information rent for the high-type consumer θ_2. Assuming the two types of consumers happen equally likely, the optimal quantity q_1 is the midpoint G of the segment IJ. As the proportion of the low type in the population increases, the monopolist finds it profitable to decrease the mark-up price to let the low type consumers purchase a higher quantity. Thus, as ρ increases, the optimal design G moves along the $\theta_1' s$ demand toward M. Hence, if the low type constitutes the whole population, it is optimal to sell as much as possible to the low type; that is, the optimal design is M. By contrast, if the low-type proportion is below a certain threshold, the monopolist's optimal design of the marginal tariff for q_1 is A, which means not serving the low type. The monopolist's attempt to balance the extraction of information rents with providing consumers with the incentives to consume according to the true preferences results in a pricing schedule that depends on the distribution of consumer heterogeneity.

In short, we have seen that compared with the price structures in the world of marginal-price rational consumers, pricing for average-price-biased consumers offers a quantity premium, efficiency at the bottom, distortion at the top (the highest type consumes less than the first-best allocation). Another difference in terms of the distribution of consuming participants between the two nonlinear pricing worlds is:

Proposition 2. *Assuming zero fixed costs, the marginal type under the nonlinear pricing scheme for average-price-biased consumers is the first-best marginal type and, of course, lower than under the pricing scheme for marginal-price rational consumers, that is,* $\overline{\theta} < \theta^*$.

Thus, pricing for average-price-responding consumers allows more people to consume than in the world of marginal-price rational consumers, favoring the distributional purpose of regulators. Moreover, because the markup in the world of pricing for average-price-biased consumers does not depend on the consumer type distribution, the marginal type does not either, even in the presence of fixed costs.

5 Welfare comparisons

Although the trade-off between efficiency and equity is less severe in the world of pricing for average-price-biased consumers, evaluating the effects on welfare is important. Given consumers exhibit average-price bias, one may think the world of nonlinear pricing for average-price-acting consumers faces a lower welfare than the world of marginal-price-acting consumers. Surprisingly, in many cases, welfare in the nonlinear pricing world with average-price-biased consumers is higher than the world of marginal-price rational consumers. One reason is that the low types are able to consume more in the world of average-price response, and this transfers to an increase in consumer surplus if the type distribution weighs heavily on low types. Another reason is that pricing for average-price-biased consumers serve more participants than the scheme for fully rational consumers. Overall, the welfare effect is ambiguous, depending on the type distribution, consumer preferences, and cost structures. I illustrate the contrast in welfare between the two regimes using a model with quadratic preferences, and consider how variations in consumer heterogeneity and cost structures lead to different welfare changes. Note that I compare the welfare between the two regimes for regulated pricing. In other words, consumer surplus between the two regulated pricing for different consumer behaviors is compared, given the regulating degrees that equalize the firms' profits across the two consuming behaviors.[6]

Consider quadratic utility

$$U(q,\theta) = \theta q - \frac{\gamma}{2}q^2 \tag{59}$$

$$C(q) = cq, \;\; \gamma > 0, \; c > 0 \tag{60}$$

Note that different type distribution leads to different price schedule. I illustrate welfare under two types of consumer heterogeneity: uniform distribution and Pareto distribution. Both distributions result in quantity premia for average-price biased consumers, but quantity discount (uniform distribution) and quantity premium (Pareto distribution) for marginal-price rational consumers.

[6] Martimort and Stole (forthcoming) compare the welfare between the two regimes under monopoly pricing and under different cost structures. I focus on the regulatory pricing and also consider the welfare comparison for a type of consumer heterogeneity other than uniform distribution. Considering a different type distribution besides the uniform distribution is important because the type distribution directly affects the tariff shapes.

5.1 Uniform distributed heterogeneity and linear costs without fixed costs

Let θ be uniformly distributed on $[\theta_0, \theta_1]$. Assume $\theta_0 < c < \theta_1$. In the standard model where consumers respond to marginal prices,

$$Q^*(\theta) = \arg\max_q u(q,\theta) - cq - r^*(\theta_1 - \theta)q = \max\left\{0, \frac{\gamma((r^*+1)\theta - r^*\theta_1 - c)}{\gamma}\right\} \tag{61}$$

and the corresponding optimal nonlinear price has marginal price

$$MP^* = -\frac{\gamma r^*}{r^*+1}q + \frac{r^*\theta_1 + c}{r^*+1} \tag{62}$$

$$P^* = -\frac{\gamma r^*}{2(r^*+1)}q^2 + \frac{r^*\theta_1 + c}{r^*+1}q. \tag{63}$$

The marginal type is $\theta^* = \frac{r^*\theta_1+c}{r^*+1} > c$. Consumer surplus for type θ is

$$CS^*(\theta) = \frac{((r^*+1)\theta - c - r^*\theta_1)^2}{2\gamma(r^*+1)}, \tag{64}$$

and ex-ante consumer surplus

$$ECS^* = \int_{\theta^*}^{\theta_1} CS^*(\theta)f(\theta)d\theta = \frac{(\theta_1 - c)^3}{6\gamma(\theta_1-\theta_0)(r^*+1)^2}. \tag{65}$$

The firm's profit and ex-ante profit are

$$\Pi^*(\theta) = \frac{[(r^*+1)\theta + c - (r^*+2)\theta_1][-(r^*+1)\theta + r^*\theta_1 + c]r^*}{2\gamma(r^*+1)} \tag{66}$$

$$E\Pi^* = \int_{\theta^*}^{\theta_1} \Pi^*(\theta)f(\theta)d\theta = \frac{r^*(\theta_1 - c)^3}{3\gamma(\theta_1-\theta_0)(r^*+1)^2}. \tag{67}$$

Now consider the average-price-biased-consumer regime; the consumption allocation is

$$\overline{Q}(\theta) = \arg\max_q qU_q(q,\theta) - cq = \max\left\{0, \frac{\theta - c}{\gamma(\overline{r}+1)}\right\}. \tag{68}$$

The price schedule and the average price rate are

$$\overline{P}(q) = \gamma\overline{r}q^2 + cq \tag{69}$$

$$\overline{AP}(q) = \gamma\overline{r}q + c. \tag{70}$$

The marginal type is $\overline{\theta} = c$. Type-specific consumer surplus, ex-ante surplus, profit, and ex ante profit are

$$\overline{CS}(\theta) = \frac{(\theta - c)^2}{2\gamma(\overline{r} + 1)^2} \tag{71}$$

$$\overline{ECS} = \int_{\overline{\theta}}^{\theta_1} \overline{CS}(\theta) f(\theta) d\theta = \frac{(\theta_1 - c)^3}{6\gamma(\theta_1 - \theta_0)(\overline{r} + 1)^2} \tag{72}$$

$$\overline{\Pi}(\theta) = \frac{\overline{r}(\theta - c)^2}{\gamma(\overline{r} + 1)^2} \tag{73}$$

$$\overline{E\Pi} = \int_{\overline{\theta}}^{\theta_1} \overline{\Pi}(\theta) f(\theta) d\theta = \frac{\overline{r}(\theta_1 - c)^3}{3\gamma(\theta_1 - \theta_0)(\overline{r} + 1)^2}. \tag{74}$$

Proposition 3. *In the quadratic preference model with constant marginal cost and zero fixed cost, for every same degree of regulation $\overline{r} = r^*$, we have*

$$\overline{E\Pi} = E\Pi^*, \overline{ECS} = ECS^*. \tag{75}$$

Although profits and consumer surpluses are equal across the regimes, note the price schedules $\overline{P}$ and P^*, optimal allocations $\overline{Q}$ and Q^*, and marginal types $\overline{\theta}$ and θ^* are very different. The features of these three parts in the average-price-biased consumer regime favors the distributional goals of the regulators. First, the price schedule for average-price-biased consumers has an increasing average rate, while the one for marginal-price-responding consumers has quantity discounts. Second, low-type consumers will consume more in the average-price-acting regime than in the marginal-price-acting regime:

$$\text{For values} \quad \theta < \frac{(\overline{r} + 1)(r^*\theta_1 + c) - c}{r^*\overline{r} + \overline{r} + r^*} \text{ , we have } \quad \overline{Q} > Q^*. \tag{76}$$

So, if pricing for biased consumers is fully regulatory, every consumer will be able to consume more than under the marginal-price response; that is, $\overline{r} = 0 \Rightarrow \overline{Q}(\theta) > Q^*(\theta)$ for all values θ.

Third, the marginal consumer type under the average-price-biased-consumer regime coincides with the first-best marginal type, $\overline{\theta} = c$, and of course below the marginal type under the second-degree price discrimination. When the prices under marginal-price-response regime become more highly regulated ($r^* \to 0$), the marginal type θ^* will fall to c, the level of the marginal type under the biased-consumer pricing regime.

Consequently, the allocation of consumption and the distribution of the marginal participants are the reasons why consumers overall do not lose surplus for the average price bias. Indeed, the population above type θ^* would earn a lower surplus under the average-price-acting regime[7]:

$$\int_{\theta^*}^{\theta_1} \overline{CS}(\theta) f(\theta) d\theta = \frac{(\theta_1 - c)^3(3r^2 + 3r + 1)}{6\gamma(\theta_1 - \theta_0)(r+1)^5} < ECS^*. \tag{77}$$

5.2 Pareto-distributed heterogeneity

Consider the situation where consumer heterogeneity is distributed according to the Pareto distribution with scale $s > 0$ and shape $\alpha > 0$. The probability density function and the cumulative distribution function are, respectively,

$$f(x) = \frac{\alpha s^\alpha}{x^{\alpha+1}} \tag{78}$$

$$F(x) = 1 - \frac{s^\alpha}{x^\alpha}. \tag{79}$$

The hazard rate $\frac{f(\theta)}{1-F(\theta)} = \frac{\alpha}{\theta}$, which is decreasing in θ. Figure 2 displays the density of the Pareto distribution under different values of shapes. We can see the Pareto distribution has the support (s, ∞) and a fat right tail. Indeed, the mean is $\frac{\alpha s}{\alpha - 1}$ for $\alpha > 1$, and the median is $s2^{1/\alpha}$. This depicts a society with mostly low-type consumers living with a rare few of extremely high types.

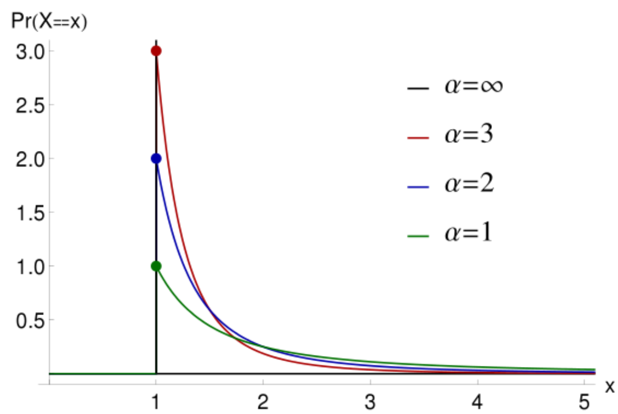


FIGURE 2: This figure depicts the Pareto density with scale $s = 1$ and different shape α.

Proposition 4. *In the quadratic preference model with constant marginal cost, zero fixed cost, and Pareto-distributed consumer heterogeneity, regulatory pricing for average-price-biased consumers improves social welfare, compared to pricing for marginal-price rational consumers:*

$$\textit{Given } \overline{r}, r^* : \overline{E\Pi} = E\Pi^*, \textit{ then } \overline{ECS} > ECS^*. \tag{80}$$

[7]To ensure profits in the two regimes equal, under regulated pricing, notice $r^* = \overline{r}$.

The reason for a welfare improvement despite average-price bias in the consumer response is that pricing for average-price-biased consumers offers features favoring the distributional goals of the regulators. For the Pareto distribution of consumer types, the price schedules in both response regimes feature quantity premia; that is, greater consumption is required to pay a higher marginal tariff:

$$P^*(q) = \frac{\gamma r^*}{\alpha - r^*}\frac{q^2}{2} + \frac{\alpha c}{\alpha - r^*}q \tag{81}$$

$$\overline{P}(q) = \gamma \overline{r} q^2 + cq \tag{82}$$

$$\overline{\theta} = c \tag{83}$$

However, the low types in the average-price regime consume more units than in the marginal-price regime, and more people are able to consume in the average-price world:

$$Q^*(\theta) = \max\left\{0, \frac{(\alpha - r^*)\theta - \alpha c}{\alpha\gamma}\right\} \tag{84}$$

$$\overline{Q}(\theta) = \max\left\{0, \frac{\theta - c}{\gamma(\overline{r}+1)}\right\} \tag{85}$$

$$\text{So,}\quad \overline{Q} > Q^* \quad \text{for} \quad \theta < \frac{\alpha c \overline{r}}{(\alpha - r^*)(\overline{r}+1) - \alpha} \tag{86}$$

$$\theta^* = \frac{\alpha c}{\alpha - r^*} > c = \overline{\theta}. \tag{87}$$

As the prices under marginal-price response regime become more highly regulated, the marginal type θ^* will fall to the marginal type level of the biased-consumer pricing regime; that is, $r^* \searrow 0 \Rightarrow \theta^* \searrow \overline{\theta}$.

The favors in consumption allocation and participation toward low-type consumers are further emphasized by the Pareto distribution that heavily weighs on the low types, leading to higher welfare in the average-price world.

Note that the fact that pricing for biased consumers expands the participating population plays a key role in enhancing social welfare, compared with pricing for rational consumers. Indeed, the monopoly pricing in the world of average-price consumers offers higher consumer surplus and higher social welfare than the other world, but the population above type θ^* earns a lower welfare under the average-price regime:

$$\int_{\theta^*}^{\infty} \overline{CS}(\theta)dF(\theta) < ECS^* \quad \text{but} \quad \overline{ECS} = \int_{\overline{\theta}}^{\theta^*} \overline{CS}dF(\theta) + \int_{\theta^*}^{\infty} \overline{E\Pi}(\theta)dF(\theta) > ECS^* \tag{88}$$

$$\int_{\theta^*}^{\infty} \overline{\Pi}(\theta)dF(\theta) > E\Pi^* \quad \text{but} \quad \overline{E\Pi} = \int_{\overline{\theta}}^{\theta^*} \overline{E\Pi}dF(\theta) + \int_{\theta^*}^{\infty} \overline{E\Pi}(\theta)dF(\theta) < E\Pi^* \tag{89}$$

$$\int_{\theta^*}^{\infty} \overline{W}(\theta)dF(\theta) < EW^* \quad \text{but} \quad \overline{EW} = \int_{\overline{\theta}}^{\theta^*} \overline{W}dF(\theta) + \int_{\theta^*}^{\infty} \overline{W}(\theta)dF(\theta) > EW^*. \tag{90}$$

5.3 Nonlinear costs

I now consider the welfare comparison in the case of nonlinear costs. For tractability, I assume consumer heterogeneity is distributed according to Pareto density.

Proposition 5. *In the quadratic preference model with Pareto-distributed consumer heterogeneity and strictly convex cost (and zero fixed cost), the regulated pricing for average-price-biased consumers offers higher welfare than the regulated pricing for marginal-price rational consumers.*

The reason for a welfare improvement is that the pass-through of the marginal cost is bigger but less salient in the case of average-price bias.[8] With the strict convexity $c_2 > 0$, an increase in c_2 by dc_2 increases the marginal cost by qdc_2. This extra marginal cost is only partly passed along to unbiased consumers in marginal prices:

$$P^* = \frac{(\gamma r^* + \alpha c_2)q^2}{2(\alpha - r^*)} + \frac{\alpha c_1 q}{\alpha - r^*} \tag{91}$$

$$\Rightarrow \frac{d}{dc_2} MP^*(q) = \frac{\alpha}{\alpha - r^*} q \ < q. \tag{92}$$

By contrast, the extra marginal cost is doubly passed along to the marginal price, but the consumer perceives only half of the pass-through:

$$\overline{P} = (\gamma \overline{r} + c_2)q^2 + c_1 q \tag{93}$$

$$\Rightarrow \frac{d}{dc_2} \overline{MP}(q) = 2q \tag{94}$$

$$\text{and} \quad \frac{d}{dc_2} \overline{AP}(q) = q. \tag{95}$$

As a result, the monopolist favors consumers with average-price bias, and not surprisingly, the welfare under regulatory pricing for biased consumers will be higher than under regulated prices for marginal-price-responding consumers. The reverse may happen when the firm has fixed costs.

Proposition 6. *In the quadratic preference model with Pareto-distributed consumer heterogeneity and a positive fixed cost (and zero marginal cost), the regulated pricing for average-price-biased consumers may offer higher welfare or lower welfare than the regulated pricing for marginal-price rational consumers, depending on the proportion of low-type consumers.*

[8] This result agrees with the finding in Martimort and Stole (forthcoming). They show that in the uniform distributed heterogeneity setting, the monopolist prefers the average-price consumer regime to the marginal-price consumer case because the pass-through of the marginal cost is also bigger but less salient in the case of average-price bias. Notice that the pass-through of the marginal cost to the price schedules for biased consumers is the same between uniform heterogeneity and Pareto heterogeneity, because the price schedule for average-price-responding consumers does not depend on the type distribution.

The loss in welfare in the bias case happens because the firm is not able to pass through any part of the fixed costs to the tariff, but to the marginal type. This inability causes a loss not only in the firm's profit, but also the consumer surplus, because the population of participants in the bias case is shrunk.

Under the conventional regime with rational consumers, part of the fixed cost is passed to consumers by an increase in the fixed tariff fee. This is not easily done for the biased consumers, because introducing a fixed fee to the price schedule will be misunderstood as an increase in the marginal price for biased consumers. Hence, none of the fixed cost is passed to biased consumers through the tariff:

$$P^*(q) = \frac{\gamma r^* q^2}{2(\alpha - r^*)} + \frac{\alpha c_0}{\alpha + r^* - 2} \tag{96}$$

$$\overline{P}(q) = \gamma \overline{r} q^2. \tag{97}$$

Alternatively, the fixed cost is passed to biased consumers through filtering participants:

$$\theta^* = \frac{\alpha\sqrt{2c_0\gamma}}{\sqrt{(\alpha - r^*)(\alpha + r^* - 2)}} \tag{98}$$

$$\overline{\theta} = \frac{(\overline{r}+1)\sqrt{\overline{r}c_0\gamma}}{\overline{r}}. \tag{99}$$

This fixed-cost pass-through increases the marginal type in the average-price-biased consumers to above the first-best level, that is $\overline{\theta} > c = 0$. Furthermore, the marginal type in the bias case may be above the second-degree level. Indeed, the difference in ex-ante profits between the two regimes is proportional to the difference in the marginal types. Under monopoly pricing, when $\alpha > 2 + \sqrt{2}$, we have $\overline{\theta} > \theta^*$ and $\overline{\Pi} < \Pi^*$, and the reverse otherwise.

Under regulatory pricing, the two participant populations in the two regimes coincide. The silence of cost-to-tariff pass-through and the loss in distributional merit in terms of the participant population tilt the welfare under average-price-biased consumers toward ambiguity compared with the total surplus under the rational-consumer regime. In the Pareto distribution of consumer types, the welfare comparison between the two regimes does not depend on the fixed-cost value but on the type-distribution parameter and the regulating degree. For example, if $\alpha = 2.5, r^* = 1, \overline{r} \approx 0.308$, we have equal profits and $\overline{ECS} > ECS^*$. If $\alpha = 3, r^* = 1, \overline{r} = 0.5$, then $\overline{\Pi} = \Pi^*$ and $\overline{ECS} = ECS^*$. However, if $\alpha = 3, r^* = 0.1, \overline{r} \approx 0.299$, then $\overline{\Pi} = \Pi^*$ and $\overline{ECS} < ECS^*$.

6 Conclusion

This paper characterizes nonlinear pricing, covering monopoly pricing and regulatory pricing, when consumers respond to average prices rather than marginal prices. I find that pricing biased consumers provides a basis for the efficiency of increasing block tariffs. Pricing in the case of average-price-biased consumers also favors distributional goals of the regulators because it offers efficiency at the bottom, distortion at the top, and the first-best participation. These features are on sharp contrast to the results in the nonlinear pricing for rational consumers.

Despite the suboptimal response of consumers, pricing in the world of average-price-biased consumers offers higher welfare than the world of rational consumers. The reason is that the markup in the biased-consumer world favors low-type consumers in a way that it allows the low types to purchase larger quantities and enables broader participation of the population. The presence of fixed costs of production or a shift in consumer heterogeneity toward high types may lead the world of nonlinear pricing for biased consumers to suffer a welfare loss.

A few extensions include considering the case in which only part of the population suffers average-price bias. This situation is consistent with the empirical evidence in De Bartolome (1995); Shaffer (2018). Empirical work that evaluates the efficiency costs of the current increasing block tariffs (IBT) used in electricity compared with the optimal tariff in the world of marginal-price rational consumers and the world of average-price-biased consumers is also valuable, especially given that we have only seen the work by Borenstein (2012) on the distortion cost of IBT away from a break-even flat-rate tariff.

References

Borenstein, S. (2009). To what electricity price do consumers respond? Residential demand elasticity under increasing-block pricing. Working Paper. April 2009.

Borenstein, S. (2012). The redistributional impact of nonlinear electricity pricing. *American Economic Journal: Economic Policy*, 4(3):56–90.

Carbajal, J. C. and Ely, J. C. (2016). A model of price discrimination under loss aversion and state-contingent reference points. *Theoretical Economics*, 11(2):455–485.

Courty, P. and Hao, L. (2000). Sequential screening. *Review of Economic Studies*, 67(4):697–717.

Dahan, M. and Strawczynski, M. (2000). Optimal income taxation: An example with a u-shaped pattern of optimal marginal tax rates: Comment. *American Economic Review*, pages 681–686.

De Bartolome, C. A. (1995). Which tax rate do people use: Average or marginal? *Journal of Public Economics*, 56(1):79–96.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, 47(2):315–372.

Diamond, P. A. (1998). Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. *American Economic Review*, pages 83–95.

Eliaz, K. and Spiegler, R. (2008). Consumer optimism and price discrimination. *Theoretical Economics*, 3(4):459–497.

Esponda, I. and Pouzo, D. (2016). Berk–Nash Equilibrium: A Framework for Modeling Agents With Misspecified Models. *Econometrica*, 84(3):1093–1130.

Farhi, E. and Gabaix, X. (2020). Optimal taxation with behavioral agents. *American Economic Review*, 110(1):298–336.

Grubb, M. D. (2009). Selling to overconfident consumers. *American Economic Review*, 99(5):1770–1807.

Hahn, J.-H., Kim, J., Kim, S.-H., and Lee, J. (2018). Price discrimination with loss averse consumers. *Economic Theory*, 65(3):681–728.

Heidhues, P. and Kőszegi, B. (2008). Competition and price variation when consumers are loss averse. *American Economic Review*, pages 1245–1268.

Ito, K. (2013). How do consumers respond to nonlinear pricing? Evidence from household water demand. Working paper.

Ito, K. (2014). Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing. *American Economic Review*, 104(2):537–63.

Liebman, J. B. and Zeckhauser, R. J. (2004). Schmeduling. Working paper. october 2004.

Martimort, D. and Stole, L. (forthcoming). Nonlinear pricing with average-price bias. *American Economic Review: Insights*.

Maskin, E. and Riley, J. (1984). Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, pages 175–208.

Mussa, M. and Rosen, S. (1978). Monopoly and product quality. *Journal of Economic Theory*, 18(2):301–317.

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, pages 180–212.

Shaffer, B. (2018). Rational, lazy or confused? evidence of misperception in consumer responsiveness to nonlinear prices. Working paper. September.

Sobel, J. (1984). Non-linear prices and price-taking behavior. *Journal of Economic Behavior & Organization*, 5(3-4):387–396.

Spiegler, R. (2012). Monopoly pricing when consumers are antagonized by unexpected price increases: A “cover version” of the Heidhues–Kőszegi–Rabin model. *Economic Theory*, 51(3):695–711.

Stole, L. A. (2007). Price discrimination and competition. *Handbook of Industrial Organization*, 3:2221–2299.

Wichelns, D. (2013). Enhancing the performance of water prices and tariff structures in achieving socially desirable outcomes. *International Journal of Water Resources Development*, 29(3):310–326.

Appendix

Omitted Proofs

Demonstration of Corollary 1. When consumers respond to marginal price and $U(q,\theta) = \theta V(q)$, the optimal pricing scheme satisfies

$$\frac{U_q - C_q}{U_q} = r^* \cdot \frac{1-F}{f} \cdot \frac{U_{q\theta}}{U_q} \tag{100}$$

$$\Rightarrow \frac{MP - MC}{MP} = \frac{r^*}{\epsilon} = \frac{r^*}{\frac{f}{1-F} \cdot \theta}. \tag{101}$$

Note $\frac{d\epsilon}{dq} = \frac{d\epsilon}{d\theta}\frac{d\theta}{dq} = (\frac{dh}{d\theta}\theta + h(\theta))\frac{d\theta}{dq}$, where $h(\theta)$ is the hazard rate $\frac{f}{1-F}$. Because $\frac{d\theta}{dq} > 0$, the increasing hazard rate implies the increasing elasticity, thus decreasing the marginal price, and decreasing the average rate.

When consumers respond to average price, the optimal pricing scheme satisfies

$$\frac{U_q - C_q}{U_q} = \overline{r} \cdot \frac{-qU_{qq}}{U_q} \tag{102}$$

$$\Rightarrow \frac{AP - MC}{AP} = \overline{r} \cdot \frac{-qU_{qq}}{U_q}. \tag{103}$$

Therefore, the increasing curvature implies the increasing average price. □

Proof of Proposition 2. Recall that the optimal allocation in the two regimes must satisfy:

$$U_q(Q^*(\theta), \theta) - C'(Q^*(\theta)) = r^* \cdot \frac{1-F(\theta)}{f(\theta)} \cdot U_{q\theta}(Q^*(\theta), \theta) \tag{104}$$

$$U_q(\overline{Q}(\theta), \theta) - C'(\overline{Q}(\theta)) = -\overline{r} \cdot \overline{Q}(\theta) \cdot U_{qq}(\overline{Q}(\theta), \theta). \tag{105}$$

Without fixed costs, the marginal types $\overline{\theta}$ and θ^* only need to satisfy $\overline{Q}(\overline{\theta}) = 0$ and $Q^*(\theta^*) = 0$, respectively.

With $\overline{Q}(\overline{\theta}) = 0$, marginal utility equals the marginal cost. In other words, the marginal type under the pricing scheme for biased consumers is the first-best marginal type level.

Now, we show $\overline{\theta} < \theta^*$. Consider the function $g(\theta; q)$ of θ with parameter q:

$$g(\theta; q) = U_q(q, \theta) - C'(q) - r^* \cdot \frac{1-F(\theta)}{f(\theta)} \cdot U_{q\theta}(q, \theta). \tag{106}$$

We have $g(\overline{\theta}; 0) = -\frac{1-F(\overline{\theta})}{f(\overline{\theta})} \cdot U_{q\theta}(0, \overline{\theta}) < 0$ and $g(\theta_1; 0) = U_q(0, \theta_1) - C'(0) > 0$. Hence, by the intermediate value theorem, the root of the function $g(\theta; 0)$ must lie between $\overline{\theta}$ and θ_1. In other words,

$\theta^* \in (\overline{\theta}, \theta_1)$. □

Proof of Proposition 4. Assume the model of quadratic utility, Pareto distribution of consumer type, and total cost $C(q) = cq$. If consumers respond to marginal prices, the consumption allocation, marginal type, price schedule, consumer surplus, and profits are

$$Q^*(\theta) = \max\left\{0, \frac{(\alpha - r^*)\theta - \alpha c}{\alpha\gamma}\right\} \tag{107}$$

$$MP^*(q) = \frac{\gamma r^* q + \alpha c}{\alpha - r^*} \tag{108}$$

$$P^*(q) = \frac{\gamma r^*}{\alpha - r^*}\frac{q^2}{2} + \frac{\alpha c}{\alpha - r^*} q \tag{109}$$

$$\theta^* = \frac{\alpha c}{\alpha - r^*} > c \tag{110}$$

$$CS^*(\theta) = \frac{[(\alpha - r^*)\theta - \alpha c]^2}{2\gamma\alpha(\alpha - r^*)} \tag{111}$$

$$ECS^* = \int_{\theta^*}^{\infty} CS^*(\theta) f(\theta) d\theta = \frac{s^\alpha c^{2-\alpha}\left(\frac{\alpha - r^*}{\alpha}\right)^{\alpha-1}}{\gamma(\alpha-1)(\alpha-2)} \tag{112}$$

$$\Pi^*(\theta) = \frac{r^*[(\alpha - r^*)\theta + \alpha c][(\alpha - r^*)\theta - \alpha c]}{2\gamma\alpha^2(\alpha - r^*)} \tag{113}$$

$$E\Pi^* = \int_{\theta^*}^{\infty} \Pi^*(\theta) f(\theta) d\theta = \frac{r^* s^\alpha c^{2-\alpha}\left(\frac{\alpha - r^*}{\alpha}\right)^{\alpha-1}}{\gamma\alpha(\alpha-2)}. \tag{114}$$

When consumers respond to average prices, the results are

$$\overline{Q}(\theta) = \frac{\theta - c}{\gamma(\overline{r} + 1)} \tag{115}$$

$$\overline{AP}(q) = \gamma\overline{r}q + c \tag{116}$$

$$\overline{P}(q) = \gamma\overline{r}q^2 + cq \tag{117}$$

$$\overline{\theta} = c \tag{118}$$

$$\overline{\Pi}(\theta) = \frac{\overline{r}(\theta - c)^2}{\gamma(\overline{r} + 1)^2} \tag{119}$$

$$\overline{E\Pi} = \int_{\overline{\theta}}^{\infty} \overline{\Pi}(\theta) f(\theta) d\theta = \frac{2\overline{r} s^\alpha c^{2-\alpha}}{\gamma(\alpha-1)(\alpha-2)(\overline{r}+1)^2} \tag{120}$$

$$\overline{CS}(\theta) = \frac{(\theta - c)^2}{2\gamma(\overline{r} + 1)^2} \tag{121}$$

$$\overline{ECS} = \int_{\overline{\theta}}^{\infty} \overline{CS}(\theta) f(\theta) d\theta = \frac{s^\alpha c^{2-\alpha}}{\gamma(\alpha-2)(\alpha-1)(\overline{r}+1)^2}. \tag{122}$$

For monopoly pricing, regulating degrees equal 1, pricing for average-price-biased consumers will offer lower profits but higher consumer surplus and higher social welfare than for marginal-price rational consumers:

$$\overline{ECS} - ECS^* = \frac{s^\alpha c^{2-\alpha}\left(4\left(\frac{\alpha-1}{\alpha}\right)^{\alpha-1} - 1\right)}{4\gamma(\alpha-1)(\alpha-2)} > 0 \tag{123}$$

$$\overline{E\Pi} - E\Pi^* = \frac{s^\alpha c^{2-\alpha}\left(2\left(\frac{\alpha-1}{\alpha}\right)^{\alpha} - 1\right)}{2\gamma(\alpha-1)(\alpha-2)} < 0 \tag{124}$$

$$\overline{EW} - EW^* = -\frac{s^\alpha c^{2-\alpha}\left[4\left(\frac{\alpha-1}{\alpha}\right)^{\alpha-1}\left(2-\frac{1}{\alpha}\right) - 3\right]}{4\gamma(\alpha-2)(\alpha-1)} > 0. \tag{125}$$

However, the participants in the marginal-price regime would suffer a lower welfare if they lived in the world of the average-price regime:

$$\int_{\theta^*}^{\infty} \overline{CS}(\theta)f(\theta)d\theta = \frac{s^\alpha(5\alpha^2-7\alpha+2)c^{2-\alpha}\alpha^{-\alpha}(\alpha-1)^{\alpha-3}}{8\gamma(\alpha-2)} < ECS^* \tag{126}$$

$$\int_{\theta^*}^{\infty} \overline{\Pi}(\theta)f(\theta)d\theta = \frac{s^\alpha(5\alpha^2-7\alpha+2)c^{2-\alpha}\alpha^{-\alpha}(\alpha-1)^{\alpha-3}}{4\gamma(\alpha-2)} > E\Pi^* \tag{127}$$

$$\int_{\theta^*}^{\infty} \overline{W}(\theta)f(\theta)d\theta = \frac{3s^\alpha c^{2-\alpha}(5\alpha-2)\left(\frac{\alpha-1}{\alpha}\right)^{\alpha}}{8\gamma(\alpha-2)(\alpha-1)^2} < EW^*. \tag{128}$$

Now, consider the regulatory pricing. The difference in profits between the average-price regime and the marginal-price regime is

$$\frac{s^\alpha}{\gamma(\alpha-2)c^{\alpha-2}}\left[\frac{2\overline{r}}{(\alpha-1)(\overline{r}+1)^2} - \frac{r^*}{\alpha}\left(1-\frac{r^*}{\alpha}\right)^{\alpha-1}\right]. \tag{129}$$

In order for the two regulatory pricing regimes in the two cases of consumer behaviors to be comparable, the regulating degrees $\overline{r}$ and r^* must be such that the profit difference is zero:

$$\frac{2\overline{r}}{(\alpha-1)(\overline{r}+1)^2} - \frac{r^*}{\alpha}\left(1-\frac{r^*}{\alpha}\right)^{\alpha-1} = 0. \tag{130}$$

Let the left-hand side be a quadratic function of $\overline{r}$, $g(\overline{r})$. Then, we have $g(0) < 0$. We also have

$$g\left(\frac{(\alpha-1)r^*}{2\alpha}\right) = \frac{r^*}{\alpha}\left[\frac{4}{\left(\left(1-\frac{1}{\alpha}\right)r^*+2\right)^2} - \left(1-\frac{r^*}{\alpha}\right)^{\alpha-1}\right] \tag{131}$$

$$\geq \frac{r^*}{\alpha}\left[\frac{4}{\left(\left(1-\frac{1}{\alpha}\right)r^*+2\right)^2} - \exp\left(-\frac{(\alpha-1)r^*}{\alpha}\right)\right] \tag{132}$$

$$> 0. \tag{133}$$

The last inequality happens because the term $\left(1-\frac{1}{\alpha}\right)r^*$ is between 0 and 1. Hence, we have that $g(\overline{r})$

is a continuous function and $g(0) \cdot g\left(\frac{(\alpha-1)r^*}{2\alpha}\right) < 0$. By the intermediate value theorem, the root of $g(\overline{r})$ lies between 0 and $\frac{(\alpha-1)r^*}{2\alpha}$. In other words, the comparable regulating degrees $\overline{r}$ and r^* satisfy $\overline{r} < \frac{(\alpha-1)r^*}{2\alpha}$.

The difference in expected consumer surplus (or expected social welfare under regulatory pricing) between the two price-response cases is

$$\overline{ECS} - ECS^* = \frac{s^\alpha (c-1)^{2-\alpha}}{(\alpha-1)(\alpha-2)}\left((\overline{r}+1)^{-2} - \left(1-\frac{r^*}{\alpha}\right)^{\alpha-1}\right) \tag{134}$$

$$= \frac{s^\alpha (c-1)^{2-\alpha}}{(\alpha-1)(\alpha-2)}(\overline{r}+1)^{-2}\left(1 - 2\cdot\frac{\overline{r}}{r^*}\cdot\frac{\alpha}{\alpha-1}\right) \quad \text{(using 130)} \tag{135}$$

$$> 0 \quad \left(\text{because } \overline{r} < \frac{(\alpha-1)r^*}{2\alpha}\right). \tag{136}$$

□

Proof of Proposition 5. Assume the model of quadratic utility, Pareto distribution of consumer type, and total cost $C(q) = c_1 q + \frac{c_2}{2}q^2$. When consumers respond to marginal prices, the results are

$$Q^*(\theta) = \max\left\{0, \frac{(\alpha - r^*)\theta - \alpha c_1}{\alpha(\gamma + c_2)}\right\} \tag{137}$$

$$\vartheta^*(q) = \frac{\alpha((\gamma+c_2)q + c_1)}{\alpha - r^*} \tag{138}$$

$$MP^*(q) = \frac{(\gamma r^* + \alpha c_2)q + \alpha c_1}{\alpha - r^*} \tag{139}$$

$$\theta^* = \frac{\alpha c_1}{\alpha - r^*} \tag{140}$$

$$P^*(q) = \frac{(\gamma r^* + \alpha c_2)q^2}{2(\alpha - r^*)} + \frac{\alpha c_1 q}{\alpha - r^*} \tag{141}$$

$$CS^* = \frac{((\alpha - r^*)\theta - \alpha c_1)^2}{2\alpha(\gamma + c_2)(\alpha - r^*)} \tag{142}$$

$$ECS^* = \frac{s^\alpha c_1^{2-\alpha}\alpha^{1-\alpha}(\alpha - r^*)^\alpha}{(\alpha - r^*)(c_2+\gamma)(\alpha-2)(\alpha-1)} \tag{143}$$

$$\Pi^* = \frac{((\alpha - r^*)\theta + \alpha c_1)((\alpha - r^*)\theta - \alpha c_1)r^*}{2(\alpha - r^*)(\gamma + c_2)\alpha^2} \tag{144}$$

$$E\Pi^* = \frac{s^\alpha r^* c_1^{2-\alpha}\alpha^{-\alpha}(\alpha - r^*)^\alpha}{(\alpha - r^*)(c_2+\gamma)(\alpha-2)}. \tag{145}$$

When consumers respond to average prices, the results are

$$\overline{Q}(\theta) = \max\left\{0, \frac{\theta - c_1}{\gamma\overline{r} + \gamma + c_2}\right\} \tag{146}$$

$$\overline{\vartheta}(q) = (\gamma\overline{r} + \gamma + c_2)q + c_1 \tag{147}$$

$$\overline{P}(q) = (\gamma\overline{r} + c_2)q^2 + c_1 q \tag{148}$$

$$\overline{\theta} = c_1 \tag{149}$$

$$\overline{CS} = \frac{\gamma(\theta - c_1)^2}{2(\gamma\overline{r} + \gamma + c_2)^2} \tag{150}$$

$$\overline{ECS} = \frac{s^\alpha \gamma c_1^{2-\alpha}}{(\gamma\overline{r} + c_2 + \gamma)^2(\alpha - 2)(\alpha - 1)} \tag{151}$$

$$\overline{\Pi} = \frac{(\theta - c_1)^2(2\gamma\overline{r} + c_2)}{2(\gamma\overline{r} + \gamma + c_2)^2} \tag{152}$$

$$\overline{E\Pi} = \frac{s^\alpha c_1^{2-\alpha}(2\gamma\overline{r} + c_2)}{(\gamma\overline{r} + c_2 + \gamma)^2(\alpha - 2)(\alpha - 1)}. \tag{153}$$

The differences in the profits and consumer surplus between the two regimes are

$$\overline{E\Pi} - E\Pi^* = \frac{s^\alpha}{c_1^{\alpha-2}(\alpha - 2)}\left[\frac{(2\gamma\overline{r} + c_2)}{(\gamma\overline{r} + \gamma + c_2)^2(\alpha - 1)} - \frac{\left(1 - \frac{r^*}{\alpha}\right)^{\alpha-1} r^*}{(\gamma + c_2)\alpha}\right] \tag{154}$$

$$\overline{ECS} - ECS^* = \frac{s^\alpha}{c_1^{\alpha-2}(\alpha - 2)}\left[\frac{\gamma}{(\gamma\overline{r} + \gamma + c_2)^2(\alpha - 1)} - \frac{(\alpha - \overline{r})^{\alpha-1}}{(\gamma + c_2)(\alpha - 1)\alpha^{\alpha-1}}\right]. \tag{155}$$

Similar to the Pareto distribution case with constant marginal cost, $c_2 = 0$, the regulating degrees $\overline{r}$ and r^* must equalize the profits between the two regimes:

$$\frac{(2\gamma\overline{r} + c_2)}{(\gamma\overline{r} + \gamma + c_2)^2(\alpha - 1)} - \frac{\left(1 - \frac{r^*}{\alpha}\right)^{\alpha-1} r^*}{(\gamma + c_2)\alpha} = 0. \tag{156}$$

Given this condition, the difference in consumer surplus is

$$\begin{aligned}\overline{ECS} - ECS^* &= \frac{s^\alpha \gamma}{c_1^{\alpha-2}(\alpha - 2)(\alpha - 1)(\gamma\overline{r} + \gamma + c_2)^2}\left[1 - 2 \cdot \frac{\overline{r}}{r^*} \cdot \frac{\alpha}{\alpha - 1} - \frac{\alpha \cdot c_2/\gamma}{r^*(\alpha - 1)}\right] && (157)\\ &> 0, && (158)\end{aligned}$$

because the quadratic equation (156) has the root $\overline{r}$ between 0 and $\frac{\gamma(\alpha-1)r^* - \alpha c_2}{2\gamma\alpha}$, making the last term in the surplus difference positive. □

Proof of Proposition 6. Assume the model of quadratic utility, Pareto distribution of consumer type, and total cost $C(q) = c_0$. (For tractability, assume zero marginal cost.) When consumers respond to marginal prices, the results are

$$
\begin{aligned}
Q^*(\theta) &= \max\left\{0, \frac{(\alpha - r^*)\theta}{\alpha\gamma}\right\} && (159)\\
\vartheta^*(q) &= \frac{\alpha\gamma q}{\alpha - r^*} && (160)\\
MP^*(q) &= \frac{\gamma r^* q}{\alpha - r^*} && (161)\\
VP^*(\theta) &= u(Q^*(\theta^*), \theta^*) - \frac{\theta^*}{\alpha}Q^*(\theta^*) - c_0 = 0 && (162)\\
\theta^* &= \frac{\alpha\sqrt{2c_0\gamma}}{\sqrt{(\alpha - r^*)(\alpha + r^* - 2)}} \quad \text{so as to virtual profit } VP^*(\theta^*) = 0 && (163)\\
P^*(q) &= u(Q^*(\theta^*), \theta^*) + \int_{Q^*(\theta^*)}^{q} u_q(x, \vartheta^*(x))dx = \frac{\gamma r^* q^2}{2(\alpha - r^*)} + \frac{c_0\alpha}{\alpha + r^* - 2} && (164)\\
CS^* &= \frac{(\alpha - r^*)(\alpha + r^* - 2)\theta^2 - 2\gamma c_0\alpha^2}{2\gamma\alpha(\alpha + r^* - 2)} && (165)\\
ECS^* &= \frac{s^\alpha[(\alpha - r^*)(\alpha + r^* - 2)]^{\alpha/2}}{2^{\alpha/2-1}\gamma^{\alpha/2}c_0^{\alpha/2-1}\alpha^{\alpha-1}(\alpha - 2)(\alpha + r^* - 2)} && (166)\\
\Pi^* &= \frac{r^*(\alpha - r^*)(\alpha + r^* - 2)\theta^2 - 2\gamma c_0\alpha^2(r^* - 2)}{2\gamma\alpha^2(\alpha + r^* - 2)} && (167)\\
E\Pi^* &= \frac{s^\alpha[(\alpha - r^*)(\alpha + r^* - 2)]^{\alpha/2}}{2^{\alpha/2-1}\gamma^{\alpha/2}c_0^{\alpha/2-1}\alpha^{\alpha}(\alpha - 2)}. && (168)
\end{aligned}
$$

When consumers respond to average prices, the results are

$$
\begin{aligned}
\overline{Q}(\theta) &= \max\left\{0, \frac{\theta}{\gamma(\overline{r} + 1)}\right\} && (169)\\
\overline{\vartheta}(q) &= q(\overline{r} + 1)\gamma && (170)\\
\overline{P}(q) &= \gamma\overline{r}q^2 && (171)\\
\overline{\Pi}(\theta) &= \frac{\overline{r}\theta^2}{\gamma(\overline{r} + 1)^2} - c_0 && (172)\\
\overline{\theta} &= \frac{(\overline{r} + 1)\sqrt{\overline{r}c_0\gamma}}{\overline{r}} \quad \text{so as to } \overline{\Pi}(\overline{\theta}) = 0 && (173)\\
\overline{CS} &= \frac{\theta^2}{2\gamma(\overline{r} + 1)^2} && (174)\\
\overline{ECS} &= \frac{s^\alpha\alpha\overline{r}^{\alpha/2-1}}{2\gamma^{\alpha/2}(\alpha - 2)c_0^{\alpha/2-1}(\overline{r} + 1)^\alpha} && (175)\\
\overline{E\Pi} &= \frac{2s^\alpha\overline{r}^{\alpha/2}}{\gamma^{\alpha/2}(\alpha - 2)c_0^{\alpha/2-1}(\overline{r} + 1)^\alpha}. && (176)
\end{aligned}
$$

The difference in profits between the two regimes is

$$\overline{\Pi} - \Pi^* = \frac{s^\alpha}{\gamma^{\alpha/2} c_0^{\alpha/2-1} \alpha^\alpha} \left[\frac{\overline{r}^{\alpha/2}}{(\overline{r}+1)^\alpha} - \frac{[(\alpha - r^*)(\alpha + r^* - 2)]^{\alpha/2}}{2^{\alpha/2}\alpha^\alpha} \right]. \tag{177}$$

The regulating degrees $\overline{r}$ and r^* equalizing the two profits must satisfy

$$\frac{\overline{r}}{(\overline{r}+1)^2} - \frac{[(\alpha - r^*)(\alpha + r^* - 2)]}{2\alpha^2} = 0. \tag{178}$$

The difference in consumer surplus is

$$\overline{ECS} - ECS^* = \frac{s^\alpha \alpha}{\gamma^{\alpha/2} c_0^{\alpha/2-1} (\alpha - 2)} \left[\frac{\overline{r}^{\alpha/2-1}}{2(\overline{r}+1)^\alpha} - \frac{[(\alpha - r^*)(\alpha + r^* - 2)]^{\alpha/2}}{2^{\alpha/2-1}\alpha^\alpha(\alpha + r^* - 2)} \right] \tag{179}$$

$$= \frac{s^\alpha \alpha}{\gamma^{\alpha/2} c_0^{\alpha/2-1} (\alpha - 2)} \cdot \frac{\overline{r}^{\alpha/2}}{(\overline{r}+1)^\alpha} \left[\frac{1}{2\overline{r}} - \frac{2}{\alpha + r^* - 2} \right] \text{ (applying 178).} \tag{180}$$

The welfare comparison between the two regimes is ambiguous in this case. Taking an example in which $\alpha = 2.5, r^* = 1, \overline{r} \approx 0.308$, we have equal profits and $\overline{ECS} > ECS^*$. If $\alpha = 3, r^* = 1, \overline{r} = 0.5$, then $\overline{\Pi} = \Pi^*$ and $\overline{ECS} = ECS^*$. However, if $\alpha = 3, r^* = 0.1, \overline{r} \approx 0.299$, then $\overline{\Pi} = \Pi^*$ and $\overline{ECS} < ECS^*$.

Although the welfare difference is ambiguous, regulatory pricing makes the marginal types between two pricing-response regimes equal; that is, $\overline{\theta} = \theta^*$ as long as $\overline{\Pi} = \Pi^*$.

Special case: monopoly pricing. Under monopoly pricing, the differences in profits and surplus between the two regimes are:

$$\overline{E\Pi} - E\Pi^* = \frac{s^\alpha \left[2^{-\alpha+1} - 2^{1-\alpha/2} \left(1 - \frac{1}{\alpha}\right)^\alpha \right]}{(\alpha - 2) c_0^{\alpha/2-1} \gamma^{\alpha/2}} \tag{181}$$

$$\overline{ECS} - ECS^* = \frac{s^\alpha \left[2^{-\alpha-1}\alpha^3 + 2^{-\alpha-1}\alpha - 2^{-\alpha}\alpha^2 - 2^{1-\alpha/2}\alpha^{-\alpha+1}(\alpha-1)^{\alpha+1} \right]}{\gamma^{\alpha/2} c_0^{\alpha/2-1} (\alpha-1)^2 (\alpha-2)} < 0. \tag{182}$$

We have that $\overline{E\Pi} < E\Pi^*$ for values $\alpha > 2 + \sqrt{2}$, and $\overline{E\Pi} > E\Pi^*$ for $2 < \alpha < 2 + \sqrt{2}$. The threshold $2 + \sqrt{2}$ in α is also the threshold for marginal type: $\overline{\theta} < \theta^*$ for values $2 < \alpha < 2 + \sqrt{2}$. However, total welfare under monopoly pricing for average-price biased consumers is always lower than for marginal-price rational consumers that is, $\overline{EW} < EW^*$, $\forall \alpha > 2$. □