Nonlinear Pricing of Food in Village Economies*

Orazio Attanasio† Elena Pastorino‡

(PRELIMINARY AND INCOMPLETE)

December 2011
Current Draft: September 2012

Abstract

We consider a nonlinear pricing model, in which consumers' marginal willingness to pay and absolute ability to pay are unobserved, to explain the nonlinearity of unit prices of basic food items (bulk discounting) in developing countries. We model consumers as budget-constrained and allow their outside option to trading with a seller to depend on their private information. We show that, under certain conditions, a nonlinear pricing problem with budget-constrained consumers is equivalent to a nonlinear pricing problem with so-called countervailing incentives, in which consumers are not budget-constrained but their reservation utility depends on their type. Hence, by applying standard techniques, we obtain a simple characterization of optimal nonlinear pricing. Based on this characterization, we show that, contrary to the prediction of standard nonlinear pricing models, nonlinear pricing can positively affect the level of consumption and the surplus enjoyed by the purchasers of the smallest quantities, typically the poorest consumers, compared to linear pricing. In particular, purchasers of small quantities can be offered larger quantities than under first best. Lastly, we prove that our model is nonparametrically identified and derive nonparametric estimators of the model primitives that can be used to assess the impact of nonlinear pricing on welfare. These estimators can be easily implemented using individual-level data commonly available for beneficiaries of conditional cash transfer programs in developing countries.

*We thank Tom Holmes, Patrick Kehoe, Kyoo-il Kim, Jean-Marc Robin, Kjetil Storesletten, and Joel Waldfogel for their comments.

†University College London and Institute for Fiscal Studies.

‡Federal Reserve Bank of Minneapolis and University of Minnesota.

1 Introduction

In this paper we propose a model of price discrimination to rationalize the pricing patterns of basic food items in rural Mexico and Colombia, in particular the occurrence of quantity discounts in the form of unit prices declining in quantity. We allow consumers to differ in their marginal willingness to pay and absolute ability to pay for a good. We explicitly model consumers' budget-constraints and allow consumers' outside options to trading with a seller to depend on their private information.

We first prove that the model can be formally thought of as an instance of a screening model with countervailing incentives, that is, a model in which consumers are not budget-constrained but their reservation utility depends on their type. Based on known techniques for these problems, we show that a simple characterization of the optimal (nonlinear) pricing contract can be obtained. By relying on this characterization, we then prove that nonlinear pricing can positively affect the level of consumption and the surplus enjoyed by the purchasers of the smallest quantities, typically the poorest consumers.

Next, we show that the model is nonparametrically identified and can be nonparametrically estimated using individual-level purchasing data, typically available for beneficiaries of conditional cash transfer programs in developing countries. The goal of the paper is to quantify the relative importance of second- versus third-degree price discrimination for the observed patterns of purchases (quantities and prices) and to measure the impact of asymmetric information, outside options, and budget constraints on consumer surplus, producer surplus, and welfare. Based on the estimated model, we can also evaluate the impact on prices and welfare of alternative pricing mechanisms, like linear pricing schemes, and food subsidization measures aimed at softening the budget constraints of the poorest households.

2 A Nonlinear Pricing Model

2.1 Environment

Consider a market populated by one producer/seller and one consumer/buyer, who enter into a contract to trade a quantity $q \in [0, \infty)$ for the monetary transfer t.[1] Benefits from trade depend

[1] This formulation is analogous to one in which there are many buyers. As discussed below, the optimal selling mechanism we derive is robust to many aspects of collusion behavior. However, the consumer may have preferences for other goods, some of which may be offered by the seller. Our formulation with budget constraints, as we discuss below, captures the fact that the consumer may have preferences for other goods not purchased from the seller. As for the case of other goods sold by the seller, we are implicitly assuming that the consumer's utility is separable in the price and quantity of each such good and that the seller prices his goods independently.

on a parameter θ privately known by the consumer and continuously distributed between $\underline{\theta}$ and $\overline{\theta}$, with cumulative distribution function $F(\theta)$ and density $f(\theta)$, positive for $\theta \in (\underline{\theta}, \overline{\theta})$. (We maintain that $\underline{\theta} > 0$.) We assume that the consumer has a fixed amount of resources available to purchase the good. We denote by $I(\theta)$ the budget of the consumer for the good, which depends on the consumer's type θ. We assume that $I(\cdot)$ is $\mathcal{C}^2$ and $I'(\theta) \geq 0$.

Upon trade, the consumer obtains utility $v(\theta, q) - t$, where $v(\cdot)$ is $\mathcal{C}^2$, $v_q(\theta, \cdot) > 0$, and $v_{qq}(\theta, \cdot) \leq 0$. Let $\overline{u}(\theta)$ indicate the consumer's utility from not purchasing the good. Our problem differs from a standard nonlinear pricing problem in two ways: (a) we allow the reservation utility of a consumer to depend on her type, as in so-called *countervailing incentives* models; and (b) we explicitly model the consumer's budget constraint. The consumer's type θ, thus, can be interpreted as an index that summarizes information about the consumer's marginal willingness to pay, absolute ability to pay, and consumption possibilities alternative to trading with the seller. We think of θ as a single index capturing potentially multiple characteristics of a consumer with appropriate weights.

The seller, who is uninformed about the value of θ, obtains profit $t - c(q)$, where $c(\cdot)$ is $\mathcal{C}^2$. We normalize the seller's reservation profit to zero. Observe that the seller's payoff is not directly affected by θ. The total surplus from the trade between the seller and the consumer is $s(\theta, q) = v(\theta, q) - c(q)$, maximal at the first-best quantity $q^{FB}(\theta)$. As usual, we assume that types can be ranked according to their marginal utility from trade, as stated in the next assumption.

(A1) $v_{\theta q}(\theta, q) > 0$ *for* q *positive.*

Timing is as follows. The seller proposes a contract to the consumer on a take-it-or-leave-it basis, where the verifiable variables are the transfer $t(\theta)$, that is, the payment for the total quantity purchased, and the quantity $q(\theta)$. The consumer is said to *participate* when she accepts the seller's offer to trade under the terms of the contract. The consumer then chooses a quantity for a corresponding transfer/quantity pair as long as the required payment does not exceed her budget and allows her to reach at least the level of utility $\overline{u}(\theta)$. Otherwise, no trade takes place. If no contract is signed, the consumer obtains her reservation utility level $\overline{u}(\theta)$, whereas the seller obtains his reservation profit of zero.[2] Let $x(\theta) \in \{0, 1\}$ denote the probability that type θ participates. Then, the consumer's utility when her type is θ is $u(\theta) = x(\theta)[v(\theta, q(\theta)) - t(\theta)] + [1 - x(\theta)]\overline{u}(\theta)$. We assume that when a given type participates, that type trades a single quantity $q(\theta)$

See Stole (2006) for a discussion of situations in which this component pricing is indeed optimal.

[2]Because of fixed costs, asset specialization, or exclusivity rules that may limit the consumer's ability to trade with other parties, the situation where no contract is enforced is not necessarily equivalent to trading $q = 0$ with no transfer. In this setup $q = 0$ should be interpreted as the limit when the contracted quantity becomes small. See also Jullien (2000) on this.

with probability one. As we will discuss, under our assumptions the restriction to deterministic contracts is without loss. We denote an allocation with exclusion by $\{u(\theta), q(\theta), x(\theta)\}$ and an allocation with full participation by $\{u(\theta), q(\theta)\}$. We say that an allocation induces *full participation* if all types participate.

By the revelation principle, a contract can be summarized by a menu $\{t(\theta), q(\theta), x(\theta)\}$ such that the consumer's best choice within the menu is to announce her type θ (incentive compatibility), the consumer receives utility at least equal to $\overline{u}(\theta)$ (individual rationality), and she is required to pay not more than $I(\theta)$ (budget constraint compatibility). Using the fact that $u(\theta) = v(\theta, q(\theta)) - t(\theta)$ when $x(\theta) = 1$, we rewrite the standard incentive compatibility constraint $v(\theta, q(\theta)) - t(\theta) \geq v(\theta, q(\theta')) - t(\theta')$ with $x(\theta) = 1$ as

$$u(\theta) \geq v(\theta, q(\theta')) - t(\theta') + u(\theta') - u(\theta') = u(\theta') + v(\theta, q(\theta')) - v(\theta', q(\theta'))$$

for all θ and θ'. The incentive compatibility and individual rationality constraints are then given by

$$\text{(IC)} \;\; u(\theta) \geq u(\theta') + v(\theta, q(\theta')) - v(\theta', q(\theta')) \text{ for any } \theta,\ \theta', \text{ if } x(\theta) = 1$$

$$\text{(IR)} \;\; u(\theta) \geq \max\{\overline{u}(\theta), \overline{u}\}, \text{ if } x(\theta) = 1.$$

We will maintain that $\overline{u}(\theta) \geq \overline{u}$, where $\overline{u}$ is an exogenous constant reservation utility level. The formulation of the consumer's participation constraint as (IR) is simply to facilitate comparisons of the properties of the optimal nonlinear pricing contract across different environments in which the consumer's reservation utility may or may not depend on her type. Analogously, by using the fact that the seller's profit from trade when $x(\theta) = 1$ can be rewritten as

$$t(\theta) - c(q(\theta)) = v(\theta, q(\theta)) - c(q(\theta)) + t(\theta) - v(\theta, q(\theta)) = s(\theta, q(\theta)) - u(\theta)$$

we can express the seller's profit from a consumer of type θ as $x(\theta)[s(\theta, q(\theta)) - u(\theta)]$. Then, the seller's optimal contract solves the program

$$\text{(P1)} \quad \max_{\{u(\theta), q(\theta), x(\theta)\}} \int_\theta x(\theta)[s(\theta, q(\theta)) - u(\theta)] f(\theta) d\theta$$

subject to

$$\text{(IC)} \;\; u(\theta) \geq u(\theta') + v(\theta, q(\theta')) - v(\theta', q(\theta')) \text{ for any } \theta,\ \theta', \text{ if } x(\theta) = 1$$

$$\text{(IR)} \;\; u(\theta) \geq \max\{\overline{u}(\theta), \overline{u}\}, \text{ if } x(\theta) = 1$$

$$\text{(BC)}\ \ I(\theta) \geq t(\theta), \text{ if } x(\theta) = 1$$

$$\text{(EX)}\ \ u(\theta) = \overline{u}(\theta), \text{ if } x(\theta) = 0.$$

The constraint (BC) captures the fact that the consumer is budget constrained. The constraint (EX) prescribes that types excluded from trade receive their reservation utility. We say that an allocation $\{u(\theta), q(\theta), x(\theta)\}$ is *implementable* if it satisfies all the above constraints. Finally, note that abstracting from transfers when trade does not take place is immaterial. If the seller paid any transfer to all excluded types, he could reduce the amount of the transfer to each such type by any small amount without affecting the consumer's behavior. We will also maintain the following assumption:

(A2) *Homogeneity.* There exists a quantity profile $\overline{q}(\theta)$ such that the allocation with full participation $\{\overline{u}(\theta), \overline{q}(\theta)\}$ is incentive compatible, that is, $\overline{u}'(\theta) = v_\theta(\theta, \overline{q}(\theta))$ and $\overline{q}(\theta)$ is increasing.

Note that assumption (A1) implies that $\overline{u}'(\theta)$ is positive. So, a sufficient condition for $\overline{u}(\theta) \geq \overline{u}$ is $\overline{u}(\underline{\theta}) \geq \overline{u}$. Moreover, since

$$\overline{u}''(\theta) = v_{\theta\theta}(\theta, \overline{q}(\theta)) + v_{\theta q}(\theta, \overline{q}(\theta)) d\overline{q}(\theta)/d\theta$$

the requirement that $\overline{q}(\theta)$ be increasing implies that $\overline{u}(\theta)$ is convex when $v_{\theta\theta}(\theta, \overline{q}(\theta)) \geq 0$. This will be the case of interest in our empirical analysis where, for reasons of model identification, we will specialize the consumer's utility function as $v(\theta, q) = \theta\nu(q)$, as common in the literature on second-degree price discrimination.

2.2 Discussion

Here we first discuss how to interpret the budget constraint (BC) in our framework, we then discuss the way our framework captures the possibility of competition among sellers, and, lastly, we discuss the likely importance of income effects in our setting.

The Budget Constraint. Consider a static utility maximization problem over two goods, good q and the numeraire c, and suppose the consumer faces a standard budget constraint on their purchase and a subsistance-level constraint on the consumption of c. We show here that this problem can be formally reinterpreted as a maximization problem over one good, in which the consumer faces a budget constraint for the purchase of that good only and this budget constraint may or may not bind at the consumer's chosen quantity of q.

Specifically, consider a consumer with quasi-linear preferences over good q and the numeraire

c. Suppose that the consumer is characterized by a parameter θ that affects her valuation for q, $u(q,\theta)$, and her budget, $Y(\theta)$. The consumer's problem is

$$\max_{q,c}\{u(q,\theta)+c\}$$

$$\text{s.t. } T(q)+c \leq Y(\theta) \text{ and } c \geq \underline{c}(\theta).$$

Since at an optimum $T(q)+c = Y(\theta)$, it follows $c = Y(\theta)-T(q)$, so the problem can be equivalently restated as

$$\max_{q}\{u(q,\theta)+Y(\theta)-T(q)\}$$

$$\text{s.t. } T(q) \leq Y(\theta)-\underline{c}(\theta).$$

Redefine the consumer's utility as $v(q,\theta)=u(q,\theta)+Y(\theta)$ and let $I(\theta)=Y(\theta)-\underline{c}(\theta)$. Then, the consumer's problem can be equivalently restated as

$$\max_{q}\{v(q,\theta)-T(q)\}$$

$$\text{s.t. } T(q) \leq I(\theta)$$

which is the case we consider. Note that the dependence of $I(\theta)$ on θ may be due to the dependence of the consumer's total income, $Y(\theta)$, on θ or of her 'required' level of consumption of the numeraire, $\underline{c}(\theta)$, on θ. In the special case in which $u(q,\theta)=\theta u(q)$ and $Y(\theta)=Y+y\theta$, it follows

$$v(q,\theta)=u(q,\theta)+Y(\theta)=\theta[u(q)+y]+Y.$$

Letting $\nu(q)=u(q)+y$, the consumer's problem can be rewritten as

$$\max_{q}\{\theta\nu(q)+Y-T(q)\}$$

$$\text{s.t. } T(q) \leq I(\theta)$$

where Y is an irrelevant constant. The case in which utility is multiplicatively separable in type and valuation for quantity is the one we will primarily focus on for reasons of model identification. We discuss this in detail below.

The Competitive Case. Suppose now that there exist multiple sellers, who behave as local monopolists in separate geographical markets. Let $T^f(q)$ denote the tariff of the seller's competitor $f \in \{1,\ldots,F\}$. By interpreting $\overline{u}(\theta)$ as $\overline{u}(\theta)=\max_f\{v(\theta,q^f(\theta))-T^f(q)\}$, that is, as the consumer's utility when purchasing the quantity $q^f(\theta)$ at price $T^f(q)$ from seller f, it is immediate that our model naturally nests the best-response problem of a firm in an oligopoly model of price competition (under vertical product differentiation, given our homogeneity assumption).

Income Effects. For simplicity we abstract from income effects. Wilson (1993; Chapter 7) shows that the optimal price schedule is quite sensitive to consumers' incomes if consumers mainly differ in their income elasticities of demand. Instead, consumers' marginal rate of substitutions and the optimal price schedule are insensitive to the optimal tariff when income elasticities are small and/or consumers' residual incomes are large in relation to their expenditures on the seller's good. In intermediate cases, income effects can influence the magnitude of the seller's profit margins but do not necessarily alter the basic form of the tariff. Allowing for the dependence of the optimal tariff on the range of customers' incomes in the presence of income effects represents an important departure from standard (linear utility) tariff design. We interpret our explicit accounting of consumers' budget constraints as a first step towards analyzing the importance of consumer's resources for optimal pricing strategies.[3]

3 A Relaxed Problem

Given primitives $v(\theta, q)$, $F(\theta)$, $c(q)$, $\overline{u}(\theta)$, and $I(\theta)$, consider an allocation $\{u(\theta), q(\theta)\}$ that is a full-participation solution to problem (P1). To characterize it, we will proceed following the approach commonly used to solve the standard nonlinear pricing problem without type-dependent reservation utilities or budget constraints. Namely, we will first consider a relaxed problem in which the constraint (BC) is dropped, problem (P2). We will then show that, under assumptions, a solution to this relaxed problem satisfies (BC) and, hence, it is also a solution to (P1). We will then consider a relaxed problem in which the constraint (IR) is replaced by the constraint $u(\theta) \geq \overline{u}$, problem (P3), and show that under conditions a solution to this relaxed problem satisfies (IR) and, hence, it is also a solution to (P1). Hence, in a formal sense the nonlinear pricing problem in which consumers have type-dependent reservation utilities and the one in which they are budget-constrained are equivalent.

Differently from the standard nonlinear pricing problem, however, in our case the relaxed problem (P2) is of interest per se. As we show here, (P2) formally captures both instances in which consumers have access to more attractive options than trading with the seller, in which case

[3] See Leslie (1998) for a welfare analysis of second- and third-degree discrimination in the pricing of tickets for Broadway shows within a structural model that accounts for income effects. In his model the quality of theater seats (the counterpart of our quantity dimension) is taken as given. Instead, we solve for the optimal price and quantity schedules and take into account changes in them when we conduct our counterfactual exercises.

their reservation utility naturally depends on θ, and situations in which a consumer's reservation utility is independent of her type but consumers are budget-constrained. Therefore, (P2) is empirically the problem of interest because, as we will show below, estimating it allows us to verify in a particular context whether: (a) the standard nonlinear pricing model applies; (b) if it does not, whether budget constraints or type-dependent reservation utilities are the source of the departure from the standard model.

Observe, as mentioned, that the relaxed problem in which (BC) is dropped is a nonlinear pricing problem with countervailing incentives in which the consumer's reservation utility depends on θ. The characterization of an optimal pricing contract for this problem is straightforward based on known techniques. We will rely on the methods developed by Jullien (2000).

3.1 Implementability of Outside Utility: A Specialization

Note that assumption (A2) requires $\overline{u}'(\theta) = v_\theta(\theta, \overline{q}(\theta))$ and $\overline{q}(\theta)$ increasing. We provide here a specialization of $\overline{u}(\theta)$ to show that $\overline{u}(\theta)$ can be easily constructed in a way that is consistent with our notion, discussed above, that $\overline{u}(\theta)$ captures alternative consumption possibilities for the consumer, one particular example being trading with other sellers.

Specifically, suppose that the reservation utility $\overline{u}(\theta) \geq \overline{u}$ is achieved by the consumer when she faces the price schedule $\overline{c}(q)$. Then, by consumer's optimality,

$$\overline{u}(\theta) = \max_q \{v(\theta, q) - \overline{c}(q)\} = v(\theta, \overline{q}(\theta)) - \overline{c}(\overline{q}(\theta)) \geq \overline{u} \tag{1}$$

where $\overline{q}(\theta) = \arg\max_q \{v(\theta, q) - \overline{c}(q)\}$ and $\overline{c}(\cdot)$ satisfies

$$\begin{cases} \overline{c}'(q) = v_q(\theta, q) \\ \overline{c}(\overline{q}(\theta)) = I(\theta). \end{cases} \tag{2}$$

Observe that the allocation with full participation $\{\overline{u}(\theta), \overline{q}(\theta)\}$ is implementable under this specialization of $\overline{u}(\theta)$. To see this, note that (IR) is satisfied by construction at the allocation $\{\overline{u}(\theta), \overline{q}(\theta)\}$. Further, note that

$$\overline{u}'(\theta) = v_\theta(\theta, \overline{q}(\theta)) + [v_q(\theta, \overline{q}(\theta)) - \overline{c}'(\overline{q}(\theta))]\, \overline{q}'(\theta) = v_\theta(\theta, \overline{q}(\theta)) \tag{3}$$

by the definition of $\overline{q}(\theta)$ and (2). Moreover, since $\overline{c}'(q) = v_q(\theta, q) > 0$ and $\overline{q}'(\theta) = I'(\theta)/\overline{c}'(\overline{q}(\theta))$, if $I'(\theta) \geq 0$ then $\overline{q}(\theta)$ is increasing. So, assumption (A2) is satisfied.

Note that it must be $\overline{u}(\underline{\theta}) = \overline{u}$. To see this, note that given all (IC)'s hold, the constraint

$v(\theta, q(\theta)) - t(\theta) \geq \overline{u}$ can be replaced by $v(\underline{\theta}, q(\underline{\theta})) - t(\underline{\theta}) \geq \overline{u}$. Hence, at the optimum, the participation constraint of the lowest type is binding. When $v(\theta, q) = \theta\nu(q)$ it follows that $\overline{u}(\underline{\theta}) = \underline{\theta}\nu(q(\underline{\theta})) - I(\underline{\theta}) = \overline{u}$, so

$$\overline{q}(\underline{\theta}) = (\nu)^{-1}\left(\frac{\overline{u} + I(\underline{\theta})}{\underline{\theta}}\right). \tag{4}$$

We provide next an example of how to construct $\overline{c}(q)$ based on $v(\theta, q)$ and $I(\theta)$, when $I(\theta)$ is a linear function of θ. (An example in which $I(\theta)$ is a polynomial function of θ is included in Appendix A.)

Example with Income Linear in Type. Suppose that $v(\theta, q) = \theta\nu(q)$ and $I(\theta) = I + i\theta$ with $i > 0$. Define the function $\overline{C}(\theta)$ as $\overline{C}(\theta) = \overline{c}(\overline{q}(\theta))$. Thus, $\overline{C}(\theta) = I + i\theta$. Let $\overline{\theta}(q)$ be the inverse mapping of $q = \overline{q}(\theta)$, which is well-defined given that $\overline{q}'(\theta) > 0$ by assumption (A2). In this case, we have

$$\begin{cases} \overline{c}'(q) = \overline{\theta}(q)\nu'(q) \\ \overline{c}(q) = I + i\overline{\theta}(q). \end{cases}$$

From the first expression in the above system it follows $\overline{\theta}(q) = \overline{c}'(q)/\nu'(q)$ and substituting this into the expression for $\overline{c}(q)$ in the above, we obtain

$$\overline{c}(q) = I + i\overline{\theta}(q) = I + i\frac{\overline{c}'(q)}{\nu'(q)} \Leftrightarrow \frac{\overline{c}'(q)}{\overline{c}(q) - I} = \frac{\nu'(q)}{i}. \tag{5}$$

Integrating both sides of (5) from $\overline{q}(\underline{\theta})$ to q, we obtain

$$\log[\overline{c}(x) - I]_{\overline{q}(\underline{\theta})}^{q} = \frac{1}{i}[\nu(x)]_{\overline{q}(\underline{\theta})}^{q} \Leftrightarrow \log[\overline{c}(q) - I] - \log[\overline{c}(\overline{q}(\underline{\theta})) - I] = \frac{1}{i}\nu(q) - \frac{1}{i}\nu(\overline{q}(\underline{\theta}))$$

or, equivalently,

$$\frac{\overline{c}(q) - I}{\overline{c}(\overline{q}(\underline{\theta})) - I} = \exp\left\{\frac{1}{i}\left[\nu(q) - \nu(\overline{q}(\underline{\theta}))\right]\right\}. \tag{6}$$

Using the fact that, by definition, $\overline{c}(\overline{q}(\underline{\theta}))- I = i\underline{\theta}$ and the expression in (4), (6) can be rewritten as

$$\overline{c}(q) = I + i\underline{\theta}\exp\left\{\frac{\nu(q) - \nu(\overline{q}(\underline{\theta}))}{i}\right\} = I + i\underline{\theta}\exp\left\{\frac{\underline{\theta}\nu(q) - I - i\underline{\theta} - \overline{u}}{i\underline{\theta}}\right\}. \tag{7}$$

Substituting this expression for $\overline{c}(q)$ into the condition $\overline{c}(q) = I + i\theta$, we also obtain

$$q = \overline{q}(\theta) = (\nu)^{-1}\left(i\log\left(\frac{\theta}{\underline{\theta}}\right) + \frac{I + i\underline{\theta} + \overline{u}}{\underline{\theta}}\right) \tag{8}$$

and

$$\theta = \overline{\theta}(q) = \underline{\theta} \exp\left\{\frac{\nu(q)}{i} - \frac{I + i\underline{\theta} + \overline{u}}{i\underline{\theta}}\right\}$$

so when $q = \overline{q}(\theta)$, we have

$$\overline{c}(\overline{q}(\theta)) = I + i\underline{\theta} \exp\left\{\frac{\underline{\theta}\nu\left((\nu)^{-1}\left(i\log\left(\frac{\theta}{\underline{\theta}}\right) + \frac{I+i\underline{\theta}+\overline{u}}{\underline{\theta}}\right)\right) - I - i\underline{\theta} - \overline{u}}{i\underline{\theta}}\right\} = I + i\theta.$$

Also,

$$\overline{u}(\theta) = \theta\nu(\overline{q}(\theta)) - I(\theta) = \theta\left(i\log\left(\frac{\theta}{\underline{\theta}}\right) + \frac{I + i\underline{\theta} + \overline{u}}{\underline{\theta}}\right) - I - i\theta$$

In the special case in which $\nu(q) = \log(q)$, $\overline{c}(q)$ becomes

$$\overline{c}(q) = I + i\underline{\theta} \exp\left\{\frac{\underline{\theta}\log(q) - \overline{u} - I - i\underline{\theta}}{i\underline{\theta}}\right\} = I + i\underline{\theta} q^{\frac{1}{i}} \exp\left\{-\frac{\overline{u} + I + i\underline{\theta}}{i\underline{\theta}}\right\}.$$

Note that if the function $q = \overline{q}(\theta)$ is known, we can alternatively proceed as follows: (1) invert the function $q = \overline{q}(\theta)$ to obtain $\overline{\theta}(q)$; (2) plug $\overline{\theta}(q)$ into the expression $\overline{c}'(q) = \overline{\theta}(q)\nu'(q)$; (3) integrate both sides of $\overline{c}'(q) = \overline{\theta}(q)\nu'(q)$ with respect to q to obtain $\overline{c}(q)$, using the boundary condition $\overline{c}(q) = I + i\overline{\theta}(q)$. Specifically, plugging $\overline{\theta}(q)$ into the expression $\overline{c}'(q) = \overline{\theta}(q)\nu'(q)$, we obtain

$$\overline{c}'(q) = \underline{\theta} \exp\left\{\frac{\nu(q)}{i} - \frac{I + i\underline{\theta} + \overline{u}}{i\underline{\theta}}\right\}\nu'(q)$$

which needs to be integrated to get $\overline{c}(q)$. In particular,

$$\overline{c}(x) = i\underline{\theta} \exp\left\{\frac{\nu(x)}{i} - \frac{I + i\underline{\theta} + \overline{u}}{i\underline{\theta}}\right\} + C$$

and using the fact that $\overline{c}(q) = I + i\theta(q)$ it follows

$$i\underline{\theta} \exp\left\{\frac{\nu(q)}{i} - \frac{I + i\underline{\theta} + \overline{u}}{i\underline{\theta}}\right\} + C = I + i\theta(q) \Rightarrow C = I$$

and

$$\overline{c}(q) = I + i\underline{\theta} \exp\left\{\frac{\nu(q)}{i} - \frac{I + i\underline{\theta} + \overline{u}}{i\underline{\theta}}\right\} = I + i\underline{\theta} \exp\left\{\frac{\underline{\theta}\nu(q) - I - i\underline{\theta} - \overline{u}}{i\underline{\theta}}\right\}$$

as obtained above. In particular, different functions $q = \overline{q}(\theta)$ will lead to different functions $\overline{c}(q)$.

Further, observe that (6) implies

$$\overline{c}'(q) = \underline{\theta} \exp\left\{\frac{1}{i}\left[\nu(q) - \nu(\overline{q}(\underline{\theta}))\right]\right\}\nu'(q)$$

so

$$\frac{\partial \overline{c}'(q)}{\partial I} = -\overline{c}'(q)\frac{1}{i}\nu'(\overline{q}(\underline{\theta}))\frac{\partial \overline{q}(\underline{\theta})}{\partial I}.$$

From $\overline{u}(\underline{\theta}) = \overline{u}$, it follows $\underline{\theta}\nu(\overline{q}(\underline{\theta})) - I - i\underline{\theta} - \overline{u} = 0$, so

$$\frac{\partial \overline{q}(\underline{\theta})}{\partial I} = -\frac{-1}{\underline{\theta}\nu'(\overline{q}(\underline{\theta}))} > 0$$

if $\underline{\theta} > 0$. Then, $\partial \overline{c}'(q)/\partial I < 0$. This result will be important for our comparative statics results below.

3.2 The Relaxed Problem Without Budget Constraints

Denote by (P2) the version of (P1) in which the constraint (BC) is dropped. So, (P2) is given by

$$\text{(P2)} \quad \max_{\{u(\theta), q(\theta), x(\theta)\}} \int_\theta x(\theta)[s(\theta, q(\theta)) - u(\theta)]f(\theta)d\theta$$

subject to

$$\text{(IC)} \;\; u(\theta) \geq u(\theta') + v(\theta, q(\theta')) - v(\theta', q(\theta')) \text{ for any } \theta,\, \theta', \text{ if } x(\theta) = 1$$

$$\text{(IR)} \;\; u(\theta) \geq \overline{u}(\theta), \text{ if } x(\theta) = 1$$

$$\text{(EX)} \;\; u(\theta) = \overline{u}(\theta), \text{ if } x(\theta) = 0$$

where we maintain that $\overline{u}(\underline{\theta}) = \overline{u}$.

Suppose that $\overline{u}(\theta) \geq v(\theta, q) - I(\theta)$ for each type. In this case it follows that if (IR) is satisfied, then (BC) is also satisfied. Assume that $q(\theta) \leq \overline{q}(\theta)$. Since $v(\theta, \cdot)$ is strictly increasing by assumption, it follows that $v(\theta, q(\theta)) \leq v(\theta, \overline{q}(\theta))$, which by (1) and (2) implies

$$\overline{u}(\theta) = \max_q\{v(\theta, q) - \overline{c}(q)\} = v(\theta, \overline{q}(\theta)) - \overline{c}(\overline{q}(\theta)) \geq v(\theta, q(\theta)) - I(\theta). \tag{9}$$

Hence, $q(\theta) \leq \overline{q}(\theta)$ is a sufficient condition for (IR) to imply (BC). When $q(\theta) \leq \overline{q}(\theta)$, a solution to (P2) is also a solution to (P1). In the next section we identify conditions under which a solution to (P2) satisfies $q(\theta) \leq \overline{q}(\theta)$ and, thus, it is also a solution to (P1).

4 Optimal Nonlinear Pricing

4.1 The Problem Without Budget Constraints

Here we characterize the solution to problem (P2), since we will focus on it. The analysis in this subsection is borrowed from Jullien (2000). Our first result is to show that, in characterizing the optimal contract for (P2), we can abstract from the possibility that the seller may find it optimal to exclude some types, say, by offering an unattractive quantity and price combination. Formally, we can prove that for any implementable allocation that entails exclusion, we can find an implementable allocation with *full participation* in which all consumer types purchase from the seller (and, naturally, viceversa). Based on this result, we can also show that full participation holds if $s(\theta, \overline{q}(\theta)) \geq \overline{u}(\theta)$ for all types, that is, if the seller makes a nonnegative profit from each type when offering $\overline{q}(\theta)$. By a slight abuse of notation, we denote an optimal contract by $\{t(\theta), q(\theta), x(\theta)\}$ and an optimal allocation by $\{u(\theta), q(\theta), x(\theta)\}$.

We formalize our first result in the next lemma. The proof of this result is in Appendix A.

Lemma 1. *Suppose (1) and (2) hold. (a) An allocation $\{u(\theta), q(\theta), x(\theta)\}$ with exclusion is implementable for (P2) if, and only if, the full participation allocation $\{u(\theta), x(\theta)q(\theta) + [1 - x(\theta)]\overline{q}(\theta)\}$ is implementable for (P2). (b) If $s(\theta, \overline{q}(\theta)) \geq \overline{u}(\theta)$, then an allocation solution to (P2) entails full participation.*

As in a standard screening problem, the objective function of the seller can be equivalently restated in terms of *virtual surplus*, the surplus adjusted to take into account the informational rents induced by incentive compatibility. Of course, the main difference with respect to the standard problem is that now the (IR) constraint need not just bind for the lowest type.

Consider an optimal allocation. Let $\gamma(\theta)$ denote the shadow value associated with a uniform (marginal) reduction of the participation level $\overline{u}(\theta)$ for all types between $\underline{\theta}$ and θ. Observe that $\gamma(\theta)$ has the properties of a cumulative distribution function: it is nonnegative, increasing, and $\gamma(\overline{\theta}) = 1$. As such, it can be represented as a unit mass distribution over the set of types $[\underline{\theta}, \overline{\theta}]$: $\gamma(\theta)$ puts a positive mass only on those types for which the participation constraint binds. Since $q(\theta)$ is continuous, $\gamma(\theta)$ can only have mass points at $\underline{\theta}$ or $\overline{\theta}$.[4] Correspondingly, $d\gamma(\theta)$ is the

[4]This distribution may be discontinuous when (IR) binds at isolated points. For example, $\gamma(\underline{\theta}) > 0$ means that (IR) binds at $\underline{\theta}$ and that there is a finite mass at $\underline{\theta}$.

shadow value of the individual rationality constraint at θ, implying $d\gamma(\theta) = 0$ when $u(\theta) > \overline{u}(\theta)$. Note that, as $\gamma(\cdot)$ is constant whenever (IR) does not bind, the support of the distribution $\gamma(\cdot)$ is included in the set of types (with no informational rent) for which $u(\theta) = \overline{u}(\theta)$. The virtual surplus in our case is given by

$$\sigma(\gamma, \theta, q) = s(\theta, q) + [(F(\theta) - \gamma)/f(\theta)]v_\theta(\theta, q)$$

whereas in the standard screening problem it is given by $\sigma(1, \theta, q)$. For the solution to the seller's problem to be well-defined, we formulate a convexity and a type-separation assumption.

(A3) *For all* $\gamma \in [0, 1]$, $\sigma(\gamma, \theta, q)$ *is strictly quasi-concave in* q.

This convexity property is verified, in particular, if $s_q(\theta, \cdot)/v_{\theta q}(\theta, \cdot)$ is decreasing with q. Under (A3) the virtual surplus is maximal at a single quantity denoted by $l(\gamma, \theta) = \arg\max_q \sigma(\gamma, \theta, q)$, which, when positive, is decreasing with γ.[5]

We now formulate a further assumption, assumption (A4), in addition to the homogeneity assumption above, in order to ensure that the optimal solution does not display bunching. In the standard model, a necessary and sufficient condition for bunching not to occur is that $l(1, \theta)$ be nondecreasing. Assumption (A4) is a generalization of this requirement. Note also that the homogeneity assumption, as stated, already requires the reservation profile $\overline{u}(\theta)$ to be convex enough. Since $\overline{u}''(\theta) = v_{\theta\theta}(\theta, \overline{q}(\theta)) + v_{\theta q}(\theta, \overline{q}(\theta))d\overline{q}(\theta)/d\theta$, the term $d\overline{q}(\theta)/d\theta$ is a measure of the curvature of $\overline{u}(\theta)$.

(A4) *For all* $\gamma \in [0, 1]$, $l(\gamma, \theta)$ *is a nondecreasing function of* θ.

Sufficient conditions for (A4) are

$$\partial \left(s_q(\theta, q)/v_{\theta q}(\theta, q)\right)/\partial\theta \geq 0 \text{ and } d\left(F(\theta)/f(\theta)\right)/d\theta \geq 0 \geq d\left([1 - F(\theta)]/f(\theta)\right)/d\theta. \qquad (10)$$

The first condition in (10) states that the marginal benefit of increasing the slope of the utility profile is weakly increasing with the type. When this condition holds, convex quantity profiles are optimal for the seller, implying that the monotonicity condition for incentive compatibility, $q(\theta)$, is easier to satisfy. The second condition in (10) states that, as the type increases, the relative weight of types above θ compared to types below θ decreases, and the seller becomes progressively more concerned about the rents left below θ. A sufficient condition for this is the

[5] The standard condition for concavity, that is, the requirement that $s(\cdot)$ be concave in q and $v_\theta(\cdot)$ convex in q, may not suffice. If $v_\theta(\cdot)$ is nonlinear, $\sigma(\gamma, \theta, q)$ cannot be concave for all distributions $F(\cdot)$ and all values of γ. A general proof of existence of an optimal contract is in Jullien (2000).

logconcavity of f.

Analogous to Jullien (2000), we consider two cases, which we term the *weakly-convex* and the *highly-convex* case, since the optimal contract differs across these two cases in the degree of convexity of the reservation utility profile, $\overline{u}(\theta)$. We start by defining Θ to be the set of interior types for which the (IR) constraint can bind,

$$\Theta = \{\theta : l(1,\theta) \leq \overline{q}(\theta) \leq l(0,\theta)\}.$$

Also, let $\overline{\gamma}(\theta)$ be defined by $\overline{q}(\theta) = l(\overline{\gamma}(\theta),\theta)$. If $\overline{q}(\theta) > 0$, then $\overline{\gamma}(\theta)$ is given by

$$\overline{\gamma}(\theta) = F(\theta) + f(\theta)s_q(\theta,\overline{q}(\theta))/v_{\theta q}(\theta,\overline{q}(\theta)) \tag{11}$$

(since $\sigma_q(\gamma,\theta,q) = 0$, $\overline{\gamma}(\theta)$ as defined implies that $\overline{q}(\theta) = l(\gamma(\theta),\theta)$ when $\gamma(\theta) = \overline{\gamma}(\theta)$).

There are circumstances in which the set Θ has a simple structure. This is the case when $\overline{q}(\cdot)$ crosses $l(\gamma,\cdot)$ for at most one type. Under such a single crossing property, it is easy to verify that Θ is an interval (possibly empty). Then, the solution to the seller's problem depends on the comparison between the slope of $\overline{q}(\theta)$ and the slope of $l(\overline{\gamma},\cdot)$. In particular, borrowing the same terminology as in Jullien (2000), we term the situation in which $\overline{q}(\cdot)$ crosses each profile at most once from above to below the weakly-convex case and the situation in which $\overline{q}(\cdot)$ crosses each profile at most once from below to above the highly-convex case. A crucial condition for the following lemma to hold is that $\overline{\gamma}(\theta)$ is increasing on Θ. Since $\overline{q}(\theta) = l(\overline{\gamma}(\theta),\theta)$ on Θ, by total differentiation it follows

$$l_\gamma(\overline{\gamma}(\theta),\theta)d\overline{\gamma}(\theta)/d\theta = d\overline{q}(\theta)/d\theta - l_\theta(\overline{\gamma}(\theta),\theta)$$

so a sufficient condition for $\overline{\gamma}(\theta)$ to be increasing on Θ is that $d\overline{q}(\theta)/d\theta \leq l_\theta(\overline{\gamma}(\theta),\theta)$ on Θ.

Lemma 2. *(Weakly-Convex Case) If $\overline{q}(\cdot)$ is continuous, Θ is nonempty, and $d\overline{q}(\theta)/d\theta \leq l_\theta(\overline{\gamma}(\theta),\theta)$ on Θ, then Θ is an interval $[\theta_1,\theta_2]$, (IR) binds on Θ, and $q(\theta) = l(0,\theta)$ if $\theta < \theta_1$, $q(\theta) = \overline{q}(\theta)$ if $\theta \in [\theta_1,\theta_2]$, and $q(\theta) = l(1,\theta)$ if $\theta > \theta_2$.*

Proof: The proof follows from the proofs of Propositions 1 and 3 and Lemma 3 in Jullien (2000). □

In the standard nonlinear pricing model, the incentive issue consists in preventing the consumer from understating her type, that is, incentive constraints are downward binding. Here the consumer may, instead, prefer to overstate her type, that is, incentive constraints can be upward binding, when the reservation utility is large for high levels of θ. Preventing an understatement is achieved by reducing the quantity (and the transfer) for smaller types: as larger types face a higher marginal benefit from trade, this makes the understatement less attractive to them. Preventing an overstatement, instead, is best achieved by increasing the quantity for larger types, taking again advantage of the fact that the consumer's marginal valuation of quantity is increasing with θ: larger and larger quantities are less and less attractive to smaller types.

Intuitively, from these observations it follows that the optimal contract in our case can exhibit *overproduction* as well as underproduction compared to the first-best level, contrary to sole underproduction implied by the standard model. Indeed, the sign of $F(\theta) - \gamma(\theta)$ tracks the direction where the incentive constraints matter more: it is negative if the downward incentive constraints dominate and positive if the upward incentive constraints dominate at θ. Therefore, the optimal contract displays overproduction when $\gamma(\theta) < F(\theta)$ and underproduction otherwise. In particular, in the weakly-convex case the participation constraint binds for a single interval of types and there is overproduction for low types ($\theta < \theta_1$) and underproduction for high types ($\theta > \theta_2$). We now turn to consider the highly-convex case.

Lemma 3. *(Highly-Convex Case) If $d\overline{q}(\theta)/d\theta \geq l_\theta(\overline{\gamma}(\theta), \theta)$ on Θ, then there exists a constant γ such that $q(\theta) = l(\gamma, \theta)$. If $0 < \gamma < 1$, then (IR) binds both at $\underline{\theta}$ and at $\overline{\theta}$.*

Proof: The proof follows from the proof of Proposition 2 in Jullien (2000). □

This result applies when the slope of $\overline{u}(\cdot)$ is high or low in which case Θ is empty. For instance, the standard nonlinear pricing model à la Baron-Myerson (1982) is one such special case: if $\overline{q}(\theta)$ is uniformly smaller than $l(1, \theta)$, then the (IR) constraint binds only at the lowest type and the quantity traded is $q(\theta) = l(1, \theta)$. The opposite case may also occur: if $\overline{q}(\theta)$ is uniformly larger than $l(0, \theta)$, then the (IR) constraint binds at the highest type and the quantity traded is $q(\theta) = l(0, \theta)$. Notice, in particular, that if $\overline{q}(\theta)$ is uniformly larger than $l(0, \theta)$, there is overproduction for all types. In all other intermediate cases in which γ is constant but different from 0 or 1, the optimal contract induces underproduction for low types and overproduction for high types.

4.2 An Equivalence Between Participation and Budget Constraints

Consider problem (P2). Here we first show that, under conditions, an optimal solution to (P2) satisfies (BC). Moreover, consumer types who are indifferent between participating and not participating, that is, for which $u(\theta) = \overline{u}(\theta)$, purchase from the seller the quantity $\overline{q}(\theta)$ at price

$I(\theta)$. In a precise sense, then, there are situations in which the participation constraint tracks consumer types whose budget constraints bind at an optimal solution.

In this Subsection we maintain that only two cases apply:

(a) either $\overline{q}(\cdot)$ is continuous, Θ is nonempty, and $d\overline{q}(\theta)/d\theta \leq l_\theta(\overline{\gamma}(\theta), \theta)$ on Θ; or

(b) $d\overline{q}(\theta)/d\theta \geq l_\theta(\overline{\gamma}(\theta), \theta)$ on Θ.

Observe that, by the derivations above, under (a) the weakly-convex case applies whereas under (b) the highly-convex case applies.

For convenience, define by $\underline{q}$ the lowest quantity purchased and by $\overline{q}$ the largest quantity purchased in a solution to (P2). Note that, since the optimal quantity schedule is increasing, it follows $\underline{q} = q(\underline{\theta})$ and $\overline{q} = q(\overline{\theta})$.

Proposition 1. *Consider problem (P2). Suppose (1) and (2) hold. Assume that $v_q(\underline{\theta}, \underline{q}) = c'(\underline{q})$, $v_q(\overline{\theta}, \overline{q}) = c'(\overline{q})$, and $\overline{q}(\theta) \geq l(1, \theta)$ for each θ. Then, $q(\theta) \leq \overline{q}(\theta)$ so when the constraint $u(\theta) \geq \overline{u}(\theta)$ is satisfied, the constraint $t(\theta) \leq I(\theta)$ is also satisfied and when $u(\theta) = \overline{u}(\theta)$, the constraint $t(\theta) = I(\theta)$ is also satisfied with $q(\theta) = \overline{q}(\theta)$. Hence, a solution to (P2) is also a solution to (P1).*

Proof: Consider a solution to (P2) and let $\theta = \theta(q)$ be the inverse of the function $q = q(\theta)$. Note that $\theta = \theta(q)$ is well-defined given that $q(\theta)$ is increasing. Define $T(q) = t(\theta(q))$. Observe first that $T'(q(\overline{\theta})) = v_q(\overline{\theta}, q(\overline{\theta}))$ by local incentive compatibility. Consider now the seller's first-order condition for the optimal choice of quantity,

$$\sigma_q(\gamma(\theta), \theta, q(\theta)) = v_q(\theta, q(\theta)) - c'(q(\theta)) - \frac{\gamma(\theta) - F(\theta)}{f(\theta)} v_{\theta q}(\theta, q(\theta)) = 0. \tag{12}$$

Since, by assumption, $v_q(\overline{\theta}, q(\overline{\theta})) = c'(q(\overline{\theta}))$, it follows that $T'(q(\overline{\theta})) = c'(q(\overline{\theta}))$. Hence, $\gamma(\overline{\theta}) = 1$. Similarly, $T'(q(\underline{\theta})) = v_q(\underline{\theta}, q(\underline{\theta})) = c'(q(\underline{\theta}))$ implies $\gamma(\underline{\theta}) = 0$ by (12). These observations together imply that the weakly-convex case applies. By the properties of the weakly-convex case, we know that $q(\theta) < \overline{q}(\theta)$ for $\theta \in [\underline{\theta}, \theta_1)$, $q(\theta) = \overline{q}(\theta)$ for $\theta \in [\theta_1, \theta_2]$, and $q(\theta) > \overline{q}(\theta)$ for $\theta \in (\theta_2, \overline{\theta}]$. However, by the proof of Proposition 3 in Jullien (2000) (ADD), if $\overline{q}(\theta) \geq l(1, \theta)$ for each θ, then $\theta \leq \theta_2$ for each θ. In particular, $\theta_2 \geq \overline{\theta}$. But since (IR) binds for all types between θ_1 and θ_2, it follows that $\overline{\gamma}(\overline{\theta}) = 1$, which implies $\theta_2 = \overline{\theta}$. Thus, $q(\theta) \leq \overline{q}(\theta)$ for all types, which also implies that $\overline{u}(\theta) \geq v(\theta, q(\theta)) - I(\theta)$, by the definition of $\overline{u}(\theta)$ and the increasingness of $v(\theta, \cdot)$. Combining $u(\theta) = v(\theta, q(\theta)) - t(\theta) \geq \overline{u}(\theta)$ and $\overline{u}(\theta) \geq v(\theta, q(\theta)) - I(\theta)$, we obtain $t(\theta) \leq I(\theta)$ as desired. Lastly, recall that $\overline{q}(\theta)$ is incentive compatible when (IR) binds by (2), so when $u(\theta) = \overline{u}(\theta)$ it follows that $q(\theta) = \overline{q}(\theta)$ and $T(q(\theta)) = \overline{c}(\overline{q}(\theta)) = I(\theta)$. $\square$

Recall that $q(\theta) \leq \overline{q}(\theta)$ for all types implies that the constraint (BC) is satisfied at an optimal solution to (P2). This is key to showing that a solution to (P2) is also a solution to (P1). Consider now a relaxed version of (P1) without (IR); we term this problem (P3). We next show that at an optimal solution to (P3) the constraint (IR) is also satisfied. Specifically, in the next proposition we show that a solution to the relaxed version of (P1) in which (IR) is dropped and substituted by the standard individual rationality constraint $u(\theta) \geq \overline{u}$ is also a solution to (P1). Hence, under assumptions, the two constraints (IR) and (BC) are equivalent.

Proposition 2. *Consider problem (P3). Suppose that $q(\theta)$ and $t(\theta)$ are strictly increasing, $t(\theta)$ is continuous, and $I'(\theta) > 0$. Then, if θ with $I(\theta) \in [t(\underline{\theta}), t(\overline{\theta})]$ there exists $\overline{u}(\theta)$ satisfying conditions (1) and (2) with $u(\theta) \geq \overline{u}(\theta)$ and $q(\theta) \leq \overline{q}(\theta)$. In particular, $u(\theta) = \overline{u}(\theta)$ and $q(\theta) = \overline{q}(\theta)$ when (BC) binds.*

Proof: Here we show that based on the primitives of (P3) and standard properties of its solution, we can construct $\overline{q}(\theta)$ and $\overline{u}(\theta)$ satisfying (1) and (2) so that the constraint $u(\theta) \geq \overline{u}(\theta)$ with $q(\theta) \leq \overline{q}(\theta)$ is satisfied at an optimal solution to (P3). Moreover, $u(\theta) = \overline{u}(\theta)$ with $q(\theta) = \overline{q}(\theta)$ when (BC) binds at an optimal solution to (P3). To start, observe that it is impossible that at an optimal solution to (P3) we have $I(\theta) < t(\underline{\theta})$ for some type θ, since the fact that $I(\theta) < t(\underline{\theta})$ would imply $I(\theta) < t(\underline{\theta}) \leq t(\theta)$ by increasingness of $t(\cdot)$. Thus, (BC) would be violated for type θ. Contradiction. Consider now a type θ for which $I(\theta) \in [t(\underline{\theta}), t(\overline{\theta})]$ at an optimal solution to (P3). There exists a type θ' such that $t(\theta') = I(\theta)$ (by the intermediate value theorem) who is offered by the seller the quantity $q(\theta') \in [q(\theta), q(\overline{\theta})]$. Using the fact that $t(\theta') = I(\theta)$, the relationship $\theta' = t^{-1}(I(\theta))$ determine such type θ' for each type θ.

Now, (BC) for type θ implies that $t(\theta) \leq I(\theta) = t(\theta')$, which, in turn, yields that $\theta \leq \theta'$, by increasingness of $t(\cdot)$, and $q(\theta) \leq q(\theta')$, by increasingness of $q(\cdot)$. Note that the incentive-compatibility constraint for type θ implies $v(\theta, q(\theta)) - t(\theta) \geq v(\theta, q(\theta')) - T(q(\theta'))$. Let

$$\begin{cases} \overline{c}'(q(t^{-1}(I(\theta)))) = v_q(\theta, q(t^{-1}(I(\theta)))) \\ \overline{c}(q(t^{-1}(I(\theta)))) = I(\theta) \end{cases}$$

where $q(\theta') = q(t^{-1}(I(\theta)))$. Specifically, construct $\overline{c}(q)$ as follows: (1) invert the function $q(\theta') = q(t^{-1}(I(\theta)))$, which corresponds to the function $q = \overline{q}(\theta)$, to obtain the function $\overline{\theta}(q)$, that is, $\theta = \overline{\theta}(q) = I^{-1}(t(q^{-1}(q(\theta'))))$; (2) plug $\overline{\theta}(q)$ into the expression $\overline{c}'(q) = v_q(\overline{\theta}(q), q)$; (3) integrate both sides of $\overline{c}'(q) = v_q(\overline{\theta}(q), q)$ with respect to q to obtain $\overline{c}(q)$, using the boundary condition $\overline{c}(q) = I + i(\overline{\theta}(q))$. (See $v(\theta, q) = \theta\nu(q)$ for a straightforward example.)

Now, recall the fact that $\overline{c}'(\overline{q}(\theta)) = v_q(\theta, \overline{q}(\theta)) > 0$, that $\overline{c}(\overline{q}(\theta)) = I(\theta)$, and that $I'(\theta) > 0$.

By applying the implicit function theorem to $\overline{c}'(\overline{q}(\theta)) - v_q(\theta, \overline{q}(\theta)) = 0$, we obtain

$$\overline{q}'(\theta) = -\frac{v_{\theta q}(\theta, q)}{v_{qq}(\theta, q) - \overline{c}''(q)}$$

whereas by applying the implicit function theorem to $\overline{c}(\overline{q}(\theta)) - I(\theta) = 0$, we obtain

$$\overline{q}'(\theta) = \frac{I'(\theta)}{\overline{c}'(\overline{q}(\theta))}.$$

These observations imply

$$-\frac{v_{\theta q}(\theta, q)}{v_{qq}(\theta, q) - \overline{c}''(q)} = \frac{I'(\theta)}{\overline{c}'(\overline{q}(\theta))} > 0$$

so $v_{qq}(\theta, q) - \overline{c}''(q) < 0$. Hence, the difference $v(\theta, q) - \overline{c}(q)$ reaches a maximum at $q(\theta')$. Combining this observation and the fact that $\overline{c}(q(\theta')) = I(\theta)$, we obtain

$$u(\theta) \geq v(\theta, q(\theta')) - T(q(\theta')) = \max_q \{v(\theta, q) - \overline{c}(q)\} = v(\theta, \overline{q}(\theta)) - I(\theta) = \overline{u}(\theta).$$

Hence, conditions (1) and (2) are satisfied with $\overline{q}(\theta) = q(\theta') \geq q(\theta)$. Repeating this argument for each θ, if follows that $q(\theta) \leq \overline{q}(\theta)$ and $u(\theta) \geq \overline{u}(\theta)$ with the desired properties. □

In the proof of Proposition 2, $\overline{q}(\theta)$ is constructed so as to map θ into $q(\theta)$ (when $t(\theta) = I(\theta)$) or $q(\theta')$ (when $t(\theta) < I(\theta)$) for each type θ with $I(\theta) \in [t(\underline{\theta}), t(\overline{\theta})]$, that is

$$\overline{q}(\theta) = \begin{cases} q(\theta), \text{ if } t(\theta) = I(\theta) \in [t(\underline{\theta}), t(\overline{\theta})] \\ q(t^{-1}(I(\theta))), \text{ if } t(\theta) < t(\theta') = I(\theta) \in [t(\underline{\theta}), t(\overline{\theta})]. \end{cases}$$

So, $\overline{c}(q) = I(\overline{\theta}(q))$ which is well-defined since, as constructed, $\overline{q}(\theta)$ is an increasing function of θ.

Observe that $q(\theta)$ and $t(\theta)$ solutions to (P3) are continuous (by the theorem of the maximum) if the relevant 'virtual surplus' function for (P3) is strictly quasi-concave and the constraint set in (P3) is convex. Increasingness of $q(\theta)$ and $t(\theta)$ solution to (P3) follows from (IC) by local incentive-compatibility.

Propositions 1 and 2 jointly imply that, under conditions, the problems (P2) and (P3) admit the same solution. This is because both (P2) and (P3) are relaxed versions of (P1) and, by Proposition 1, a solution to (P2) is also a solution to (P1) and, by Proposition 2, a solution to (P3) is also a solution to (P1). In this precise sense, under the joint conditions of Propositions 1 and 2, problems (P2) and (P3) are equivalent. (Recall that in the standard problem, this logic applies as follows: first, a solution to a relaxed version of the problem, in which the monotonicity condition for incentive compatibility is dropped, is determined. Next, it is proved that such solution also satisfies the monotonicity condition.)

Based on these results, in the rest of the paper we will focus on the solution to (P2). Then, in our empirical analysis we will verify whether the conditions of Propositions 1 (and 2) are satisfied, so that a solution to (P2) is also a solution to (P1) and, hence, (P3).

4.3 A Basis for Empirical Analysis

Consider then problem (P2). Observe that the optimal quantity schedule $q(\theta) = l(\gamma(\theta), \theta)$ satisfies

$$\sigma_q(\gamma(\theta), \theta, q(\theta)) = v_q(\theta, q(\theta)) - c'(q(\theta)) - \frac{\gamma(\theta) - F(\theta)}{f(\theta)} v_{\theta q}(\theta, q(\theta)) = 0. \tag{13}$$

In the data we typically do not observe $t(\theta)$ but, instead, a price for each offered quantity, which we denote by $T(q)$. The schedule $T(q)$ is obtained from $t(\theta)$ by substituting q for θ based on the inverse relationship $\theta = \theta(q)$, with $\theta(q) \equiv (q)^{-1}(q)$, derived from the optimal quantity schedule $q = q(\theta)$. Note that this inverse function, relating *observed* quantities to *unobserved* types, is well-defined since $q(\theta)$ is increasing in θ by incentive compatibility. This relationship will prove central to the identification of the model. Hence, we can rewrite (13) as

$$\frac{T'(q(\theta)) - c'(q(\theta))}{T'(q(\theta))} = \frac{\gamma(\theta) - F(\theta)}{f(\theta)} \frac{v_{\theta q}(\theta, q(\theta))}{v_q(\theta, q(\theta))} \tag{14}$$

since $T'(q(\theta)) = v_q(\theta, q(\theta))$. Consider the special case, standard in the literature, in which $v(\theta, q) = \theta\nu(q)$. We will focus on this specification of utility in our empirical analysis. Since in this case $T'(q) = \theta\nu'(q)$ and $v_{\theta q}(\theta, q(\theta)) = \nu'(q)$, necessary (and sufficient) conditions for an optimal solution are

$$\frac{T'(q(\theta)) - c'(q(\theta))}{T'(q(\theta))} = \frac{\gamma(\theta) - F(\theta)}{\theta f(\theta)} \tag{15}$$

and

$$T'(q(\theta)) = \theta\nu'(q(\theta)). \tag{16}$$

Equations (15) and (16) will form the basis of our identification and estimation strategy. We formulate a further assumption.

(A5) $\overline{q}(\cdot)$ crosses $l(\gamma, \cdot)$ for at most one type.

In the following we maintain that only the weakly-convex or the highly-convex case applies.

5 Nonlinear Pricing and Consumer Surplus

Consider the complete information version of our problem in which consumers are not budget-constrained and their reservation utility is independent of their type. Suppose for simplicity that the seller's cost function entails no fixed cost and a constant marginal cost, c. If the seller was allowed to fully (third-degree) price discriminate, surplus would be maximal at the first-best quantity $q^{FB}(\theta)$ but fully appropriated by the seller. On the other hand, if the seller behaved perfectly competitively and offered each consumer type her first-best quantity $q^{FB}(\theta)$ at marginal cost per unit, the first-best linear price, consumer surplus would be maximal and producer surplus zero.

Consider next the incomplete information environment of problem (P2) but assume that the consumer's reservation utility is independent of her type. This is the standard nonlinear pricing problem with $\gamma(\theta) = 1$ for all types, since in this case $\overline{u}(\theta) = \overline{u}$ for each θ and the participation constraint binds only for the lowest type. Recall that the optimal price schedule in this case is nonlinear and implies that the consumption of all types, apart from the highest one, is distorted downward compared to first best. That is, the optimal contract trades off the allocative distortion imposed to all consumers types, apart from the highest one, against the informational rent granted to all consumer types, apart from the lowest one, to induce them to reveal their type. As a result, the optimal contract prescribes underconsumption and levels of utility for all types, except for the highest one, below first best.

Consider now the incomplete information environment in which consumers' reservation utility depends on their type. We have shown that in this case overproduction, compared to first best, occurs when $\gamma(\theta) < F(\theta)$. If so, then consumer types who are offered quantities larger than first best pay marginal prices below the seller's marginal cost. In any such case, the interplay between price discrimination and incomplete information has a non-obvious impact on consumer (and producer) surplus. The question we address here is under which conditions, if any, in this case some consumers are better off when they face nonlinear prices than when they face the linear monopoly price, under complete or incomplete information, or even the linear first-best price c.

5.1 The Impact of Participation Constraints on Consumer Surplus

To this purpose, consider the weakly-convex case of our model. Note first that the linear monopoly price under complete or incomplete information is higher than the marginal price that a consumer of type θ faces in our setting when $\gamma(\theta) = 0$. This result can be easily seen from the fact that when $\gamma(\theta) = 0$ the optimal contract implies $T'(q(\theta)) < c$ for all $\theta \in [\underline{\theta}, \theta_1]$. Suppose that $v_q(\theta, \cdot)$ is strictly decreasing (it is decreasing by assumption). Recall that $T(\cdot)$ is increasing by incentive compatibility, $q^{FB}(\theta)$ solves $v_q(\theta, q) = c$, and $q(\theta)$ solves $v_q(\theta, q) = T'(q)$. These observations imply that $q(\theta) \geq q^{FB}(\theta)$ for all $\theta \in [\underline{\theta}, \theta_1]$.

Now, denote by $q_m(\theta)$ the quantity purchased and by p_m the price paid by a consumer of type θ in the linear monopoly pricing solution under complete or incomplete information. Obviously, the solution to the monopolist's pricing problem typically differs across the complete and incomplete information cases. However, the distinction between the two cases is immaterial for the point we make here. Indeed, we will only rely on the fact that $q^{FB}(\theta) \geq q_m(\theta)$, regardless of whether $q_m(\theta)$ refers to the quantity demanded by a consumer of type θ when p_m is the linear monopoly price under complete or incomplete information. For this reason, we simply refer to p_m and $q_m(\theta)$, respectively, as the monopolist's linear price and the quantity demanded by type θ.

Note that by local incentive compatibility, $v_q(\theta, q(\theta)) = T'(q(\theta)) = t'(\theta)/q'(\theta)$ and

$$u'(\theta) = v_\theta(\theta, q(\theta)) + v_q(\theta, q(\theta))q'(\theta) - t'(\theta) = v_\theta(\theta, q(\theta)).$$

So, $u(\theta) = u(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v_\theta(x, q(x))dx$. A consumer of type θ is better-off under our nonlinear pricing solution than under the linear monopoly pricing solution if, and only if,

$$u(\theta) = u(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v_\theta(x, q(x))dx > u_m(\theta) \equiv v(\theta, q_m(\theta)) - p_m q_m(\theta).$$

Let $\widetilde{\theta}$ be such that $\widetilde{\theta} \leq \theta_1$. By the mean value theorem, there exists $\theta' \in (\underline{\theta}, \widetilde{\theta})$ such that

$$\int_{\underline{\theta}}^{\widetilde{\theta}} v_\theta(x, q(x))dx = (\widetilde{\theta} - \underline{\theta})v_\theta(\theta', q(\theta')). \tag{17}$$

Now, let $v(\theta, q) = \theta\nu(q)$. As argued, for $\theta \in [\underline{\theta}, \theta_1]$, we have $q(\theta) \geq q^{FB}(\theta) \geq q_m(\theta)$. Suppose $\underline{\theta}$ is sufficiently small that $\widetilde{\theta} - \underline{\theta} > \theta'$. Hence, $v_\theta(\theta', q(\theta')) = \nu(q(\theta'))$. Recall that $\overline{u}(\theta) \geq \overline{u}$. Then, by (17)

$$u(\theta') \geq (\widetilde{\theta} - \underline{\theta})\nu(q(\theta')) > \theta'\nu(q(\theta')) \geq \theta'\nu(q_m(\theta')) > u_m(\theta') = \theta'\nu(q_m(\theta')) - p_m q_m(\theta'). \tag{18}$$

Then, $u(\theta') > u_m(\theta')$. By continuity, there exists an interval of types close to θ' for which consumer surplus is higher under nonlinear monopoly pricing than under linear monopoly pricing.

Proposition 3. *Suppose $v(\theta, q) = \theta\nu(q)$, $\overline{u}(\underline{\theta}) \geq 0$, and that the weakly-convex case applies. If $\underline{\theta}$ is sufficiently small, then under nonlinear pricing, consumers of the smallest quantities enjoy higher utility than under linear monopoly pricing and purchase quantities higher than under first-best linear pricing.*

This result has important implications for the distributional impact of nonlinear pricing. It implies that, when consumers' reservation utility is type-dependent or consumers are budget-constrained, consumers of the smallest quantities are better off under nonlinear pricing than under linear pricing. In particular, they may be able to achieve levels of utility close to first best despite the presence of private information, in situations in which this would *not* be the case if they were *not* budget-constrained. Hence, the presence of budget constraints fundamentally alters the characteristics of an optimal contract in ways that can be beneficial to the poor, if the poor are the purchasers of the smallest quantities.

6 Testable Implications

Consider the interpretation of the participation constraint as a budget constraint. In this case, the model has implications for the relationship between properties of the distributions of consumer budgets and types, on the one hand, and properties of the observed price schedule in a market (village), on the other hand.

Specifically, below we address the following questions: What is the relationship between properties of the income distribution in a village, interpreted as a proxy of demand conditions, and properties of the price schedule? Do different distributions of income call for different intensities of price discrimination?

We also explore how changes in $F(\theta)$ affect the seller's price schedule. In anticipation of the empirical analysis, we will maintain that $v(\theta, q) = \theta\nu(q)$ and that marginal cost is constant.

6.1 Changes in the Distribution of Consumer Budgets

Recall the seller's first-order condition in (15). In the standard nonlinear pricing problem, the multiplier $\gamma(\theta)$ is equal to one for each type, so changes in consumers' budgets do not affect the optimal price schedule. In our case, instead, $\gamma(\theta)$ is in general different from one. Hence, the fact that consumers are budget-constrained affects the slope of the optimal price schedule at *each* quantity, in particular at quantities purchased by consumer types whose budget constraints do *not* bind.

By inspection of (15), it follows that increases in $\gamma(\theta)$ are associated with increases in $T'(q)$. Further, from the local incentive-compatibility condition $\theta\nu'(q) = T'(q)$, the fact that $\nu'(\cdot)$ is decreasing implies that increases in $T'(q)$ are associated with decreases in q for each θ. So, increases in $\gamma(\theta)$ are associated with increases in $T'(q)$ and a downward shift in the schedule $q(\theta)$.

Now suppose that $I(\theta) = I + i(\theta)$ for some positive function $i(\theta)$ with $i'(\theta) > 0$. Consider an increase in I that results in the new budget schedule $I_\Delta(\theta)$ with $I_\Delta(\theta) \geq I(\theta)$ for all types. Denote by $\gamma_{\Delta I}(\theta)$ the new schedule of multipliers associated with the consumer's budget constraint after this increase in $I(\theta)$. Also, denote by $T_{\Delta I}(q)$ the new price schedule, by $T'_{\Delta I}(q)$ the new marginal price schedule, and by $q_{\Delta I}(\theta)$ the new quantity schedule associated with $\gamma_{\Delta I}(\theta)$. Since, by construction, $\overline{c}(\overline{q}(\theta)) = I + i(\theta)$, it follows that

$$\partial\overline{q}(\theta)/\partial I = I/\overline{c}'(\overline{q}(\theta)) > 0.$$

In turn, the fact that $\overline{u}'(\theta) = \nu(\overline{q}(\theta))$ implies that increases in $\overline{q}(\theta)$ are associated with increases in $\overline{u}'(\theta)$ for each θ. Since, with $\overline{u}(\underline{\theta}) = \overline{u}$, we can express $\overline{u}(\theta)$ as

$$\overline{u}(\theta) = \overline{u} + \int_{\underline{\theta}}^{\theta} \overline{u}'(x)dx.$$

So, increases in $I(\theta)$ achieved through increases in I are associated with increases in $\overline{u}(\theta)$.

Lemma 4. *Assume that $v(\theta, q) = \theta\nu(q)$, marginal cost is constant, and $I(\theta) = I+i(\theta)$, $i'(\theta) > 0$. Marginal increases in I are associated with marginal increases in $\overline{q}(\theta)$ and $\overline{u}(\theta)$ for each type.*

Observe that the important step in this argument is the fact that the numerator of $\partial\overline{q}(\theta)/\partial I$ is positive. Note that any positive function $i(\cdot)$ can, without loss, be thought of as a polynomial $i(\theta) = i_1\theta + \cdots + i_n\theta^n$. Suppose that $i_j \geq 0$, $1 \leq j \leq n$. Then, with $\underline{\theta} > 0$ it follows that an increase in any i_j would result in an increase in $\overline{q}(\theta)$ and, thus, in $\overline{u}(\theta)$.

Consider first the weakly-convex case. Suppose I increases in a way that is still compatible with the weakly-convex case. (For this it is enough that $\theta\nu''(\overline{q}(\theta)) - \overline{c}''(\overline{q}(\theta))$ is sufficiently large; see Appendix A for details.) By the properties of the weakly-convex case, at a new optimal allocation $\gamma_{\Delta I}(\theta)$ is constant at zero for $\theta \in [\underline{\theta}, \theta'_1)$, it is equal to $[F(\theta) - F(\theta'_1)]/[F(\theta'_2) - F(\theta'_1)]$ for $\theta \in [\theta'_1, \theta'_2]$, and it is constant at one for $\theta \in (\theta'_2, \overline{\theta}]$. Observe that increases in $\overline{q}(\theta)$ are associated with decreases in $\overline{\gamma}(\theta)$, since from (11)

$$\frac{\partial\overline{\gamma}(\theta)}{\partial\overline{q}(\theta)} = \frac{c\theta^2 f(\theta)\nu''(\overline{q}(\theta))}{[\theta\nu'(\overline{q}(\theta))]^2} \leq 0 \tag{19}$$

given that $\nu''(\cdot) \leq 0$. Thus, increases in $I(\theta)$ are associated with decreases in $\overline{\gamma}(\theta)$. In turn, this immediately implies that $\theta_1 \leq \theta_1'$ and $\theta_2 \leq \theta_2'$, which yields that $\gamma_{\Delta I}(\theta) \leq \gamma(\theta)$. So, $\gamma_{\Delta I}(\theta)$ first-order stochastically dominates $\gamma(\theta)$. In turn, from $\gamma_{\Delta I}(\theta) \leq \gamma(\theta)$ it follows that $T'_{\Delta I}(q) \leq T'(q)$ by (15) and, as argued, $q_{\Delta I}(\theta) \geq q(\theta)$. Further, the fact that $q_{\Delta I}(\theta) \geq q(\theta)$ for each θ implies that the support of purchased quantity shifts to the right after the increase in I.

Consider now the highly-convex case. In analogy to the above, we focus on increases in $I(\theta)$ through I that are still compatible with the highly-convex case. By Lemma 4, increases in I are associated with increases in $\overline{q}(\theta)$. This implies that for a type whose budget constraint binds, the purchased quantity, $q(\theta)$, increases. Since γ is constant for all θ in the highly-convex case and $l(\gamma, \theta)$ is nonincreasing in γ, it follows that γ decreases. Now, starting from $\gamma = 1$, as γ decreases it follows that the budget constraint first binds either for the lowest type only ($\gamma = 1$) or for the lowest and highest types only ($\gamma \in (0, 1)$) and eventually binds for all types ($\gamma = 0$). Also, as in the weakly-convex case, a decrease in γ causes the entire marginal price schedule to shift downward.

Lemma 5. *Assume that $v(\theta, q) = \theta\nu(q)$, marginal cost is constant, and $I(\theta) = I + i(\theta)$, $i'(\theta) > 0$. In the standard nonlinear pricing model, changes in consumer budgets do not affect the price schedule. In the presence of budget constraints, a first-order stochastic dominant improvement in the distribution of consumer budgets implies a lower marginal price of the good at each offered quantity and a right-shift in the support of purchased quantities.*

Lastly, we relate changes in consumers' budgets to changes in the intensity of price discrimination. In order to so, we start by relating changes in $T'(q)$ to changes in the intensity of price discrimination. Observe that with $T(q) = p(q)q$, under linear pricing the price per unit $T(q)/q = p(q) = p$ and the marginal price $T'(q) = p$ are constant in quantity. Under nonlinear pricing, instead, the price per unit $T(q)/q = p(q)$ and the marginal price $T'(q) = p'(q)q + p(q)$ are typically not constant in quantity. Thus, (the absolute value of) $T''(q)$ provides a natural measure of the intensity of price discrimination. (See Appendix A for a formal discussion of the relationship between the curvature of the price schedule and the extent to which the price schedule entails 'bulk discounting', that is, unit prices declining in quantity.)

To understand how changes in $I(\theta)$ affect $T''(q)$, we proceed as follows. Recall that from $T'(q(\theta)) = \theta\nu'(q(\theta))$ and the monotonicity of $\nu(\cdot)$, it follows that both in the weakly-convex and in the highly-convex cases, if $T'(q)$ shifts downward (respectively, upward), then the optimal quantity schedule $q(\theta)$ shifts upward (respectively, downward). Denote by $G(q)$ the cumulative distribution function of *observed* quantity purchases and by $g(q)$ the corresponding probability density function of quantity purchases associated with the price schedule $T(q)$. Similarly, denote by $G_{\Delta I}(q)$ the cumulative distribution function of *observed* quantity purchases and by $g_{\Delta I}(q)$ the corresponding probability density function associated with the price schedule $T_{\Delta I}(q)$. Recall that, by incentive-compatibility, the optimal quantity schedule $q(\theta)$ defines a mapping between unobserved types and observed purchased quantities, $q = q(\theta)$. Since the function $q(\cdot)$ is increasing, we can define the inverse function $\theta = \theta(q)$, where $\theta(q) = q^{-1}(q)$, that relates observed quantities to unobserved types. Note that

$$G(q) = \Pr(\tilde{q} \leq q) = \Pr(\tilde{\theta} \leq q^{-1}(q) = \theta) = F(\theta)$$

so $\theta'(q) = g(q)/f(\theta)$. From $T'(q) = \theta(q)\nu'(q)$ it follows that $T'(q) > 0$, since $\nu'(\cdot) > 0$ by assumption, and

$$T''(q) = \theta'(q)\nu'(q) + \theta(q)\nu''(q).$$

By dividing the left-hand side and the right-hand sides of the above equality by $T'(q) = \theta(q)\nu'(q)$ and using the fact that $\theta'(q) = g(q)/f(\theta)$, we obtain

$$\frac{T''(q)}{T'(q)} = \frac{\theta'(q)\nu'(q)}{T'(q)} + \frac{\theta(q)\nu''(q)}{T'(q)} = \frac{\theta'(q)}{\theta(q)} + \frac{\nu''(q)}{\nu'(q)} = \frac{g(q)}{\theta f(\theta)} + \frac{\nu''(q)}{\nu'(q)} \tag{20}$$

where $T''(q)/T'(q) = \partial \log(T'(q))/\partial q$. Note that $-\nu''(q)/\nu'(q)$ is the Arrow-Pratt measure of absolute risk aversion of the consumer. Compare $\partial \log(T'(q))/\partial q$ and $\partial \log(T'_{\Delta I}(q))/\partial q$, where

$$\frac{\partial \log(T'_{\Delta I}(q))}{\partial q} = \frac{g_{\Delta I}(q_{\Delta I})}{\theta f(\theta)} + \frac{\nu''(q_{\Delta I})}{\nu'(q_{\Delta I})}.$$

Suppose that the base utility function $\nu(\cdot)$ displays decreasing or constant absolute risk aversion, that is, $-\nu''(q')/\nu'(q') \leq -\nu''(q)/\nu'(q)$ for any $q' > q$. Assume that both $T''(q)$ and $T''_{\Delta I}(q)$ are negative, that is, the two price schedules imply quantity discounts. Then, $T_{\Delta I}(q)$ implies a greater intensity of price discrimination than $T(q)$ if, and only if,

$$\frac{\partial \log(T'_{\Delta I}(q))}{\partial q} \leq \frac{\partial \log(T'(q))}{\partial q} \Leftrightarrow \frac{g_{\Delta I}(q_{\Delta I})}{\theta f(\theta)} - \frac{g(q)}{\theta f(\theta)} \leq -\frac{\nu''(q_{\Delta I})}{\nu'(q_{\Delta I})} + \frac{\nu''(q)}{\nu'(q)}. \tag{21}$$

Recall that $T'_{\Delta I}(q) \leq T'(q)$ implies $q(\theta) \leq q_{\Delta I}(\theta)$, so $-\nu''(q_{\Delta I})/\nu'(q_{\Delta I}) \leq -\nu''(q)/\nu'(q)$. Note that, by construction, the two schedules $q(\theta)$ and $q_{\Delta I}(\theta)$ are such that $G(q) = F(\theta)$ and $G_{\Delta I}(q_{\Delta I}) = F(\theta)$. Compare now the two densities $g_{\Delta I}(q_{\Delta I})$ and $g(q)$ of quantity purchases at each pair of quantities $(q, q_{\Delta I})$ such that $G(q) = G_{\Delta I}(q_{\Delta I})$, that is, at quantities in the support of $G_{\Delta I}(q_{\Delta I})$ such that $q_{\Delta I} = G^{-1}_{\Delta I}(F(\theta))$ and at quantities in the support of $G(q)$ such that

$q = G^{-1}(F(\theta))$. By (21), if $g_{\Delta I}(q_{\Delta I})$ is sufficiently smaller than $g(q)$ at each such pair of quantities, then $T_{\Delta I}(q)$ implies a greater intensity of price discrimination than $T(q)$.[6] If both $T''(q)$ and $T''_{\Delta I}(q)$ are positive, that is, the two price schedules imply quantity premia, then

$$\frac{\partial \log(T'(q))}{\partial q} \leq \frac{\partial \log(T'_{\Delta I}(q))}{\partial q} \Leftrightarrow -\frac{\nu''(q_{\Delta I})}{\nu'(q_{\Delta I})} + \frac{\nu''(q)}{\nu'(q)} \leq \frac{g_{\Delta I}(q_{\Delta I})}{\theta f(\theta)} - \frac{g(q)}{\theta f(\theta)} \tag{22}$$

so if $g_{\Delta I}(q_{\Delta I}) \geq g(q)$ at each pair of quantities $(q, q_{\Delta I})$ selected as described, then $T_{\Delta I}(q)$ implies a greater intensity of price discrimination than $T(q)$.

Finally, suppose the base utility function $\nu(\cdot)$ displays increasing or constant absolute risk aversion. Hence, $q(\theta) \leq q_{\Delta I}(\theta)$ implies $-\nu''(q)/\nu'(q) \leq -\nu''(q_{\Delta I})/\nu'(q_{\Delta I})$. Now, if both $T''(q)$ and $T''_{\Delta I}(q)$ are negative, then by (21)

$$\frac{\partial \log(T'_{\Delta I}(q))}{\partial q} \leq \frac{\partial \log(T'(q))}{\partial q} \Leftrightarrow -\frac{\nu''(q)}{\nu'(q)} + \frac{\nu''(q_{\Delta I})}{\nu'(q_{\Delta I})} \leq \frac{g(q)}{\theta f(\theta)} - \frac{g_{\Delta I}(q_{\Delta I})}{\theta f(\theta)}$$

which is satisfied if $g(q) \geq g_{\Delta I}(q_{\Delta I})$ at each pair of quantities $(q, q_{\Delta I})$ selected as described. Instead, if $T''(q)$ and $T''_{\Delta I}(q)$ are positive, then by (22)

$$\frac{\partial \log(T'(q))}{\partial q} \leq \frac{\partial \log(T'_{\Delta I}(q))}{\partial q} \Leftrightarrow \frac{g(q)}{\theta f(\theta)} - \frac{g_{\Delta I}(q_{\Delta I})}{\theta f(\theta)} \leq -\frac{\nu''(q)}{\nu'(q)} + \frac{\nu''(q_{\Delta I})}{\nu'(q_{\Delta I})}$$

which is satisfied if $g(q)$ is sufficiently smaller than $g_{\Delta I}(q_{\Delta I})$ at each pair of quantities $(q, q_{\Delta I})$ selected as described.

We summarize these results in the following Lemmas. We first consider the case in which the utility function displays constant or decreasing absolute risk aversion.

Lemma 6. *Assume marginal cost is constant and $I(\theta) = I + i(\theta)$, $i'(\theta) > 0$. Suppose $v(\theta, q) = \theta\nu(q)$ and $\nu(q)$ displays constant or decreasing absolute risk aversion. Suppose $T'(q) > 0$ and $T'_{\Delta I}(q) > 0$. Consider a first-order stochastic dominant improvement in the distribution of consumer budgets.*

(a) If $T(q)$ and $T_{\Delta I}(q)$ imply quantity discounts and $T_{\Delta I}(q)$ is associated with a large enough decrease in the density of quantity purchases, $T_{\Delta I}(q)$ implies a greater intensity of price discrimination.

(b) If $T(q)$ and $T_{\Delta I}(q)$ imply quantity premia and $T_{\Delta I}(q)$ is associated with an increase in the density of quantity purchases, $T_{\Delta I}(q)$ implies a greater intensity of price discrimination.

[6] Since $g(q) = f(\theta)/q'(\theta)$, improvements in $I(\theta)$ lead to an increase in $g(q)$ for all q if increases in $q(\theta)$ are associated with decreases in $q'(\theta)$.

We now consider the case in which the utility function displays constant or increasing absolute risk aversion.

Lemma 7. *Assume marginal cost is constant and $I(\theta) = I + i(\theta)$, $i'(\theta) > 0$. Suppose $v(\theta, q) = \theta\nu(q)$ and $\nu(q)$ displays constant or increasing absolute risk aversion. Suppose $T'(q) > 0$ and $T'_{\Delta I}(q) > 0$. Consider a first-order stochastic dominant improvement in the distribution of consumer budgets.*
(a) If $T(q)$ and $T_{\Delta I}(q)$ imply quantity discounts and $T_{\Delta I}(q)$ is associated with a decrease in the density of quantity purchases, $T_{\Delta I}(q)$ implies a greater intensity of price discrimination.
(b) In $T(q)$ and $T_{\Delta I}(q)$ imply quantity premia and $T_{\Delta I}(q)$ is associated with a large enough increase in the density of quantity purchases, $T_{\Delta I}(q)$ implies a greater intensity of price discrimination.

6.2 Changes in the Distribution of Types

Consider now the impact of changes in the type distribution. We maintain that the support of $F(\theta)$ does not change in this comparative statics exercise, only its probability density function. Because of the interaction between $\gamma(\theta)$ and $f(\theta)$ in the seller's first-order condition (15), our model leads to different comparative statics results, compared to the standard nonlinear pricing model in which $\gamma(\theta) = 1$ for all types, when the type distribution changes.

Formally, denote by $T_F(q)$ the price schedule and by $G_F(q)$ the associated cumulative distribution function of quantity purchases under the type distribution $F(\theta)$. Similarly, denote by $T_H(q)$ the price schedule and by $G_H(q)$ the associated cumulative distribution function of quantity purchases under the type distribution $H(\theta)$, a distribution with the same support as $F(\theta)$ but different probability density function, which we denote by $h(\theta)$. Note that $F(\theta) = G_F(q)$ and $H(\theta) = G_H(q)$. Let $\lambda_F(\theta)$ and $\lambda_H(\theta)$ denote the hazard rates of $F(\theta)$ and $H(\theta)$, respectively.

Recall that a distribution $F(\theta)$ is said to dominate a distribution $H(\theta)$ according to the *hazard rate order* if $f(\theta)/[1 - F(\theta)] \leq h(\theta)/[1 - H(\theta)]$ for each θ. Note that hazard rate dominance is a weaker ordering criterion than likelihood ratio dominance: in fact, it is implied by likelihood ratio dominance. Moreover, if $F(\theta)$ dominates $H(\theta)$ in the likelihood ratio or hazard rate order, then $F(\theta)$ also first-order (and, thus, second-order) stochastically dominates $H(\theta)$. Examples of parametric families of distributions that can be ordered according to monotone likelihood ratio dominance, and, thus, hazard rate dominance, are the exponential, the binomial, the Poisson, and the normal distribution. Lastly, note that since $q(\theta)$ is increasing by incentive compatibility, if $F(\theta)$ dominates $H(\theta)$ in the likelihood ratio order (or hazard rate order), then $G_F(q)$ dominates

$G_H(q)$ in the first-order stochastic dominance sense.[7]

Suppose that the seller's marginal cost is constant. Consider the standard nonlinear pricing model. By (15), under this model $T'_F(q) \geq T'_H(q)$ at each q if, and only if,

$$\lambda_F^{-1}(\theta) = \frac{1 - F(\theta)}{f(\theta)} \geq \lambda_H^{-1}(\theta) = \frac{1 - H(\theta)}{h(\theta)}$$

at each θ, that is, if, and only if, $F(\theta)$ dominates $H(\theta)$ according to the hazard rate order. Moreover, since $F(\theta) = G_F(q)$ and $H(\theta) = G_H(q)$, it also follows that $G_F(q)$ dominates $G_H(q)$ according to first-order stochastic dominance.

Consider now our model. Since hazard rate dominance implies first-order stochastic dominance, it follows that if $F(\theta)$ dominates $H(\theta)$ according to the hazard rate order we also have $F(\theta) \leq H(\theta)$. In turn, (apart from the trivial case) the fact that $F(\theta)$ first-order stochastically dominates $H(\theta)$ implies that $h(\theta) > f(\theta)$ at low (enough) types and $h(\theta) < f(\theta)$ at high (enough) types. Denote by $\gamma_F(\theta)$ and $\gamma_H(\theta)$, respectively, the multipliers under $F(\theta)$ and $H(\theta)$, respectively. By (15) we have

$$\frac{T'_F(q) - c}{T'_F(q)} = \frac{\gamma_F(\theta) - F(\theta)}{\theta f(\theta)} = \frac{\lambda_F^{-1}(\theta)}{\theta} - \frac{1 - \gamma_F(\theta)}{\theta f(\theta)}$$

and

$$\frac{T'_H(q) - c}{T'_H(q)} = \frac{\gamma_H(\theta) - H(\theta)}{\theta h(\theta)} = \frac{\lambda_H^{-1}(\theta)}{\theta} - \frac{1 - \gamma_H(\theta)}{\theta h(\theta)}$$

so in our model $T'_F(q) \geq T'_H(q)$ if, and only if,

$$\lambda_F^{-1}(\theta) - \lambda_H^{-1}(\theta) = \frac{1 - F(\theta)}{f(\theta)} - \frac{1 - H(\theta)}{h(\theta)} \geq \frac{1 - \gamma_F(\theta)}{f(\theta)} - \frac{1 - \gamma_H(\theta)}{h(\theta)}. \tag{23}$$

Suppose $F(\theta)$ dominates $H(\theta)$ in the hazard rate order and that the weakly-convex case applies both under $F(\theta)$ and $H(\theta)$. Then, $\gamma_F(\overline{\theta}) = \gamma_H(\overline{\theta}) = 1$, which implies that (23) is certainly satisfied in a neighborhood of $\overline{\theta}$. At types θ in $(\underline{\theta}, \theta_1)$, instead, if $F(\theta)$ and $H(\theta)$ are sufficiently close to each other, then

$$\frac{1 - F(\theta)}{f(\theta)} - \frac{1 - H(\theta)}{h(\theta)} = \frac{h(\theta)[1 - F(\theta)] - f(\theta)[1 - H(\theta)]}{f(\theta)h(\theta)} \leq \frac{h(\theta) - f(\theta)}{f(\theta)h(\theta)} = \frac{1}{f(\theta)} - \frac{1}{h(\theta)} \tag{24}$$

at low enough θ's, with strict inequality at some small $\theta > \underline{\theta}$. Therefore, it can be $T'_F(q) < T'_H(q)$

[7] See Müller and Stoyan (2002) for a standard reference on stochastic orders and their relationships.

at some q even if $\lambda_F^{-1}(\theta) \geq \lambda_H^{-1}(\theta)$.

Suppose now that the highly-convex case applies both under $F(\theta)$ and $H(\theta)$. Note that if $1-\gamma_F < 1-\gamma_H$, then at $\underline{\theta}$ we have

$$\lambda_F^{-1}(\underline{\theta}) - \lambda_H^{-1}(\underline{\theta}) = \frac{h(\underline{\theta}) - f(\underline{\theta})}{f(\underline{\theta})h(\underline{\theta})} \geq \frac{h(\underline{\theta})(1-\gamma_F) - f(\underline{\theta})(1-\gamma_H)}{f(\underline{\theta})h(\underline{\theta})} = \frac{1-\gamma_F}{f(\underline{\theta})} - \frac{1-\gamma_H}{h(\underline{\theta})}$$

so $T_F'(q) \geq T_H'(q)$ at the lowest quantity. Assume now that $1-\gamma_F \geq 1-\gamma_H$, so

$$(1-\gamma_H)\left[\frac{h(\theta) - f(\theta)}{f(\theta)h(\theta)}\right] \leq \frac{h(\theta)(1-\gamma_F) - f(\theta)(1-\gamma_H)}{f(\theta)h(\theta)}.$$

In this case, if γ_H is sufficiently small and $F(\theta)$ and $H(\theta)$ are sufficiently close to each other, then at low θ's we have

$$\frac{h(\theta)[1-F(\theta)] - f(\theta)[1-H(\theta)]}{f(\theta)h(\theta)} \leq (1-\gamma_H)\left[\frac{h(\theta) - f(\theta)}{f(\theta)h(\theta)}\right] \tag{25}$$

with strict inequality at some small $\theta > \underline{\theta}$ if γ_H is close enough to zero. By the same logic as in (24), this argument implies that it can be $T_F'(q) < T_H'(q)$ at some q even if $\lambda_F^{-1}(\theta) \geq \lambda_H^{-1}(\theta)$.

Lemma 8. *Assume marginal cost is constant. Suppose that the distribution $F(\theta)$ dominates the distribution $H(\theta)$ according to the hazard rate order. Then, $G_F(q)$ first-order stochastically dominates $G_H(q)$. Then, under the standard nonlinear pricing model, $T_F'(q)$ is larger than $T_H'(q)$. Under our model, $T_F'(q)$ can be larger or smaller than $T_H'(q)$ at small quantities.*

Contrary to the standard model, then, in our model an improvement (in the likelihood ratio sense) of the type distribution and, thus, of the distribution of quantity purchases does not necessarily imply an increase in the slope of the price schedule.

7 Data

In this section, we describe our data and present some descriptive evidence based on them. We use data from the surveys collected to evaluate the Conditional Cash Transfer (CCT) programs that were implemented in Mexico, PROGRESA/*Oportunidades*, and in Colombia, *Familias en Acción*. For this reason, our data is not representative of rural Mexico or Colombia. Our justification for using these data sources is the richness of the surveys and the fact that they constitute a longitudinal data base that covers, in the case of Mexico, a long time period. We will now discuss the cases of the two countries in turn.

Mexico: PROGRESA/Oportunidades

Mexico is a middle income country that in the last few decades has witnessed many changes, including a strong process of urbanization. Its rural population, however, remains an important fraction of the total and one that is disproportionately affected by poverty. The rural poor are also perceived to be particularly vulnerable to external shocks, such as the increases in food prices and their distinctive nonlinearity in quantity. Since 1997, the Mexican government has used a Conditional Cash Transfer program as its main strategy to fight poverty in rural areas. The program was started under the Zedillo administration and was first known as PROGRESA. The program consolidated and replaced many pre-existing programs, including some consisting of in-kind transfers. PROGRESA was expanded since its inception in 1997 to cover about 50,000 localities in 2000 and has become the largest welfare program of the Mexican government effectively covering almost all of rural Mexico. Two types of rural localities were not covered by PROGRESA. The first set is constituted of localities that were perceived as not being 'marginalized' and 'poor enough' to need this type of intervention. Some of these localities were included in subsequent expansions of the program. The second set of localities were places that did not have access to enough health and education infrastructure to allow the beneficiary households to comply with the conditionalities imposed by the program at a reasonable cost. These localities were covered by other programs. These details are important for us because we use data from the evaluation of the rural component of PROGRESA/*Oportunidades* and therefore only include localities targeted by this program. In principle it would be important to perform a similar exercise for the very marginal localities not covered by PROGRESA.

Unlike many other programs targeted at rural areas, PROGRESA survived the 2000 change of administration and only changed in name, becoming *Oportunidades* under the Fox administration. In 2002, the program was expanded to some urban areas, excluding only the largest cities. While the urban program is identical to the rural one in terms of the size of the grants provided, the registration mechanism is very different. In this paper we do not study households covered by the program in urban areas, partly because they constitute a very different reality which should probably be the object of a different study and partly because the data available do not include the same detail of information on prices and unit values that is available in the rural survey.

In order to evaluate PROGRESA and estimate its impact, several large surveys were conducted in some rural areas. The so-called 'ENCEL' surveys collect extensive information on all households living in 506 rural localities from seven Mexican states. (Because of the way PROGRESA and then *Oportunidades* are targeted, the program performs a census of all the households living in all rural localities where the program operates. This survey, labeled ENCASEH, contains information on several variables, but not on consumption and expenditure.) These 506 villages were randomly allocated to two groups of 320 and 186. In the first group of localities, PROGRESA began operation in mid 1998, while in the rest of the sample the program started to operate at the end of 1999. Data were collected in March and October 1998, in May and November 1999, and in April and November 2000 in villages from both groups.

In October 2003 an additional survey was collected, which included the 506 localities and an additional 150 localities in which *Oportunidades* was still not operating in 2003. This last subset of localities belong to the set of localities that were not 'poor enough' to be included in the first expansion of PROGRESA. They are therefore systematically different from those in the original sample. This is one of the reasons why we concentrate on the (May) 1999 survey. Finally, a large fraction of the localities in the 2003 ENCEL sample were visited and households re-contacted in October 2007 for a follow-up survey. On this last occasion, localities in Chiapas were added to the sample.

The ENCEL surveys constitute the bulk of the data we use for our analysis. Our sample consists of the May 1999 survey. As mentioned above, we excluded data from the urban program roll-out to ensure a relatively homogenous sample and because price data were not available for those surveys. For this reason we also exclude the additional localities surveyed in the 2003 and 2007 ENCEL survey and use only the original 506 localities originally in the 1998 ENCEL. As we mentioned, all households living in a locality are interviewed. A large fraction of them is constituted by households who are eligible for PROGRESA/*Oportunidades*. On average, 78% of the households in our localities are beneficiaries. There is a substantial amount of variability in eligibility rates across localities. Obviously, beneficiary households are 'poorer' than non-beneficiary ones. This can be verified in a variety of dimensions, from the ownership of durables to the fraction of total consumption devoted to food. In any case, the large majority of the households in our sample are quite poor. On average, for instance, food accounts for nearly 70% of their total budget.

The ENCEL surveys contain information on a number of demographic and socio-economic variables. In what follows we make use of the following variables: the sex, age and ethnicity of the head of household, the household size, the number of children and the earnings of the household head. The locality-level average of head-earnings is calculated and used as an instrument for total food consumption. One of the main reasons for our use of the ENCEL survey is the richness of the consumption and expenditure data. In the case of food and drink (alcoholic and non alcoholic), the survey contains information on weekly expenditure and quantity purchased, for

36 goods, together with the quantity consumed and home produced. The foods included and their share in the budget can be found in Table 1 (ADD), and cover fruits and vegetables, grains and pulses, meat and animal products, and other foods. The list is supposed to be exhaustive of the foods consumed by these households. Table 2 (ADD) reports, for the same foods, the percentage of households consuming the food during the survey week.

In addition to information on food consumed and/or purchased in the last week, the survey contains information about several other items, over different time intervals. In the case of some items, such as utilities, questions are asked about expenditure in the last month, while for some others, such as clothing and furniture, the questions in the survey refer to the last six months. On all these items, however, there is only information on values spent, not on quantities. One, therefore, cannot compute unit values, as in the case of food.

Given that the survey contains information on quantities purchased and consumed as well as the value of expenditure, it is in theory possible to observe prices, or to be more precise, unit values. There are, however, a number of measurement issues. They are discussed in detail in Attanasio, Di Maro, Lechene, and Phillips (2009) (CONTINUE).

Colombia: Familias en Acción. This program, inspired by PROGRESA in Mexico, consists mainly of conditional cash transfers meant to improve the accumulation of human capital among the poorest households in rural Colombia. Beneficiaries are the poorest 20% of households living in towns with less than 100,000 inhabitants and with enough health and education infrastructure. The sample includes towns where the program started to operate in 2002 and towns where the program does not operate because they did not satisfy some of the conditions for its operation (like the presence of a bank). The data were collected from a total of 122 municipalities in rural parts of Colombia between June and October 2002.

Our sample includes clusters of households in the urban center of town as well as in rural areas (which, in turn, are divided between rural and dispersed rural, the latter being more isolated). The towns in our sample are relatively small: the median population is 20,300. Our towns are also varied: the smallest town has just over 1,000 people, while the largest has just over 120,000 inhabitants. At the household level, the sample consists of families that are potential beneficiaries of the program—that is, households with children from the poorest sectors of society. The household survey contains 11,497 households. Data were collected at both the household and the individual level (CONTINUE).

8 Empirical Analysis

The focus of this section is to provide an empirical assessment of the importance of asymmetric information and limited participation for observed pricing and purchasing patterns, as well as for consumer and producer surplus, in the context of rural markets for basic food items, specifically rice. One goal of our empirical exercise is to determine whether poor consumers, who tend to be purchasers of the smallest quantities (see Attanasio, Di Maro, Lechene, and Phillips (2009)), benefit from nonlinear pricing.

To this purpose, we first prove that model primitives are nonparametrically identified. Second, we show that nonparametric estimators can be easily obtained based on our identification results. Our strategy for the recovery of model primitives is based on Guerre, Perrigne, and Vuong (2000) (henceforth, GPV) and Perrigne and Vuong (2007) (henceforth, PV). Third, we compare nonparametric estimates, obtained from these estimators, from the standard nonlinear pricing model and from our model, when it applies, and their distributional implications.

We assume that in the data we observe the outcomes of the optimal pricing decisions on the part of sellers and of the optimal participation and purchasing choices on the part of consumers, which we characterized in previous sections. We assume that exclusion of consumers occurs only when inefficient ($s(\theta, \overline{q}(\theta)) < \overline{u}(\theta)$). Hence, without loss, we maintain that the full participation solution characterized above applies to the set of participating consumers (see Jullien (2000) for details).

8.1 Preliminaries

Note first that in the data we do not observe $t(\theta)$ but, instead, $T(q)$. In the following we treat the price schedule $T(q)$ as observed. The reason is that our data provide information about unit prices from stores in different geographical areas. We also have available unit values constructed based on extensive information on quantity purchased and expenditure on food items at the household level (see Attanasio (2006), Attanasio, Di Maro, Lechene, and Phillips (2009) for details). Thus, denoting the observed price per unit or unit value for quantity q by $p(q)$, we can compute the price schedule $T(q)$ as $T(q) = p(q)q$. The primitives defining the model structure are $v(\cdot)$, $F(\cdot)$, $c(\cdot)$, and $\overline{u}(\cdot)$. In recovering them, we maintain the following assumptions:

(IA1): The seller makes non-negative profits from each consumer's type.

(IA2): Observed prices and quantities correspond to the full participation solution of (P2), either under the weakly-convex or the highly-convex case.

(IA3): The consumer's utility function is $v(\theta, q) = \theta\nu(q)$.

(IA4): The seller's cost function entails a constant marginal cost.

(IA5): $\underline{\theta} > 0$.

By Lemma 1, assumption (IA1) is an immediate consequence of seller optimality and the fact that the consumer's private information does not directly affect the seller's profit. Assumption (IA2) defines the case under consideration. As discussed above, we can assess whether the participation constraint stands in for a budget constraint by comparing the schedules $q(\theta)$ and $\overline{q}(\theta)$. We maintain (IA3) for consistency with most of the literature on nonlinear pricing and for reasons of model identification. The motivation for (IA4) is that, as we will argue, the cost function appears in the seller's first-order condition, equation (15), only through its derivative. Since (15) provides our estimating equation for $c(\cdot)$, we can at most identify $c'(\cdot)$. Finally, again for reasons of identification, we normalize $\underline{\theta}$. We choose a positive value for convenience.

The relationship between observed quantities and unobserved (to the seller and the econometrician) types implied by the model will prove central to our empirical strategy. The fact that $q(\theta)$ is increasing (by incentive compatibility) guarantees the existence of a unique monotone mapping between a consumer's type and that type's purchased quantity (this observation is due to Maskin and Riley (1984)). As in GPV and PV, let $G(q)$ denote the fraction of consumers purchasing at most quantity q, $G(q) = \Pr(\tilde{q} \leq q)$ with density function $g(q)$. From $q = q(\theta)$ it follows $\theta = q^{-1}(q)$ and

$$G(q) = \Pr(\tilde{q} \leq q) = \Pr(\tilde{\theta} \leq q^{-1}(q) = \theta) = F(\theta)$$

which implies $g(q) = f(\theta)/q'(\theta) = f(\theta)\theta'(q)$. Intuitively, a scale normalization is necessary because two observationally equivalent structures can be obtained by transforming type θ into type $\tilde{\theta} = \lambda\theta$, for any $\lambda > 0$, and the unknown base marginal utility function $\nu'(\cdot)$ into $\tilde{\nu}'(\cdot) = \nu'(\cdot)/\lambda$ (see also PV on this). Let also $\underline{q} \equiv q(\underline{\theta})$ and $\overline{q} \equiv q(\overline{\theta})$, where $\underline{q}$ and $\overline{q}$ denote, respectively, the smallest and largest observed quantity purchased in a market. In our case a market will correspond to a village (a municipality) in Mexico (for the moment being).

8.2 Identification

Here we provide conditions for the nonparametric identification of the model primitives, $\nu(\cdot)$, $F(\cdot)$, c, and $\overline{u}(\cdot)$ (with the caveat that $\overline{u}(\cdot)$, and, hence, $I(\cdot)$ when the problem corresponds to one in which consumers are budget-constrained, are only identified for types for which the relevant constraint binds).[8] We start by discussing how to determine the type of pricing contract that applies to a particular market based solely on information on observed prices and quantities. Following the derivations above, we distinguish two cases, the weakly-convex case and the highly-convex case.

To start, observe that the seller's first-order condition (15) can be rewritten as

$$\frac{T'(q)-c}{T'(q)} = \frac{\gamma(\theta)-F(\theta)}{\theta f(\theta)} = \frac{\gamma(\theta)-G(q)}{\theta g(q)q'(\theta)} \tag{26}$$

where the last equality follows from the fact that $F(\theta) = G(q)$ and $f(\theta) = g(q)q'(\theta)$. Recall from Lemma 2 that in the weakly-convex case $\gamma(\underline{\theta}) = 0$ and $\gamma(\overline{\theta}) = 1$, so $T'(\underline{q}) - c = T'(\overline{q}) - c = 0$ since $F(\underline{\theta}) = 0$ and $F(\overline{\theta}) = 1$. Hence, $T'(\underline{q}) = T'(\overline{q})$, which can easily be tested in the data. If the restriction $T'(\underline{q}) = T'(\overline{q})$ is rejected, then the highly-convex applies. We detail below how $\gamma(\theta) = \gamma$ can be recovered from $T'(q)$, $G(q)$, and $g(q)$.

Lemma 9. *If $T'(\underline{q}) = T'(\overline{q})$, then observed prices and quantities correspond to the weakly-convex case. Otherwise, observed prices and quantities correspond to the highly-convex case.*

Recall that the standard model is a special instance of the highly-convex case in which $\gamma = 1$. Now, (26) can be alternatively rewritten as

$$G(q) = \gamma - \theta q'(\theta)[T'(q)-c]\frac{g(q)}{T'(q)} = \gamma - \frac{\theta(q)}{\theta'(q)}g(q) + c\frac{\theta(q)}{\theta'(q)}\frac{g(q)}{T'(q)}$$

so, since a one-to-one relationship exists between q and θ, we can specify an additive nonparametric regression model in which the partial regression functions $f_1(\cdot)$ and $f_2(\cdot)$ are estimated from the data

$$G(q) = \gamma + f_1\left(g(q)\right) + f_2\left(\frac{g(q)}{T'(q)}\right) \tag{27}$$

where $\theta q'(\theta) > 0$. Suppose we regress $G(q)$ on a constant, $g(q)$, and $g(q)/T'(q)$. Then, the estimated constant provides an estimate of γ. Key to this argument for the recovery of γ is that q can be treated as exogenous given that all variables and functions in (27) are deterministic. If so, no correlation between any included parameter or variable and the error term can arise (assuming measurement error is negligible). If the estimated value of γ is significantly different from one, we can reject the hypothesis that the standard nonlinear pricing model applies.

[8] Note that if we interpret $I(\theta)$ as the difference between the consumer's income and expenditure on all other goods, consistent with our formulation of the two-good utility maximization problem leading to the constraint $t(\theta) \leq I(\theta)$, then $I(\theta(q)) = \Upsilon(q)$. $\Upsilon(q)$ can be nonparametrically identified and estimated from our data on household expenditure.

Lemma 10. *If the constant term in a regression of $G(q)$ on $g(q)$ and $g(q)/T'(q)$ as in (27) is significantly different from one, then the standard nonlinear pricing model does not apply.*

Note that our approach below in the highly-convex case also allows us to identify and estimate the primitives of the standard nonlinear pricing model, $\nu(\cdot)$, $F(\cdot)$, and c, nonparametrically.

8.2.1 Weakly-Convex Case

Recall that the weakly-convex case includes both the case in which the constraint (IR) binds for one interval of interior types, $[\theta_1, \theta_2]$, and the cases in which it binds for one interval of types including the lowest one, that is, $\theta_1 = \underline{\theta}$, or one interval of types including the highest one, that is, $\theta_2 = \overline{\theta}$. Moreover, under our single-crossing assumption between $\overline{q}(\theta)$ and $l(\gamma, \theta)$, assumption (A5), such an interval is nonempty and unique. Let $q_1 = q(\theta_1)$ and $q_2 = q(\theta_2)$, where $q(\theta)$ denotes the optimal quantity schedule. Recall also that in the weakly-convex case, $\gamma(\theta) = 0$ for $q \in [\underline{q}, q_1)$ and $\gamma(\theta) = 1$ for $q \in (q_2, \overline{q}]$ (with obvious modifications if $\underline{q} = q_1$ or $q_2 = \overline{q}$), so $T'(\underline{q}) = T'(\overline{q}) = c$.

Lemma 11. *In the weakly-convex case, the marginal cost c is identified by the slope of the price schedule at the smallest or largest quantity purchased in the market considered.*

Over the interval of quantities $[q_1, q_2]$ at which (IR) binds, $\gamma(\theta)$ coincides with the truncated type distribution over $[q_1, q_2]$, so $\gamma(\theta) = [F(\theta) - F(\theta_1)] \,/\, [F(\theta_2) - F(\theta_1)]$ or, equivalently,

$$\gamma(\theta(q)) = [G(q) - G(q_1)] \,/\, [G(q_2) - G(q_1)] \,. \tag{28}$$

Note that $\gamma(\theta(q_1)) = 0$ and $\gamma(\theta(q_2)) = 1$. Hence, $\gamma(\theta(q_1)) \leq G(q_1)$ and $\gamma(\theta(q_2)) \geq G(q_2)$, with strict inequalities if $q_1 > \underline{q}$ and $q_2 < \overline{q}$. Also, the difference

$$\gamma(\theta(q)) - G(q) = \frac{G(q)\,\{1 - [G(q_2) - G(q_1)]\} - G(q_1)}{G(q_2) - G(q_1)}$$

is strictly increasing, since its slope, $g(q)\,\{1 - [G(q_2) - G(q_1)]\}\,/[G(q_2) - G(q_1)]$, is strictly positive. This is because if $[\theta_1, \theta_2]$ is nonempty, then $G(q_2) - G(q_1) > 0$ and, since it cannot simultaneously be that $q_1 = \underline{q}$ and $q_2 = \overline{q}$ without violating the assumptions of the case ((IR) cannot bind for all types in the weakly-convex case), then $G(q_2) - G(q_1) < 1$. Also, $\gamma(\theta)$ is continuous at θ_1 and θ_2. (If not, then $\gamma(\theta)$ would have a mass point at either θ_1 or θ_2 or both. But since the optimal quantity is continuous, $\gamma(\theta)$ can only have mass points at $\theta = \underline{\theta}$ or $\theta = \overline{\theta}$.) Therefore, $[T'(q) - c]/T'(q)$ equals zero also at a unique interior quantity $\widetilde{q}$, which can be computed as the quantity at which $\gamma(\theta(q))$ equals $G(q)$ or, equivalently,

$$\frac{G(\widetilde{q})\left\{1-[G(q_2)-G(q_1)]\right\}-G(q_1)}{G(q_2)-G(q_1)}=0.$$

This implies

$$G(q_2)=1-G(q_1)\left[\frac{1}{G(\widetilde{q})}-1\right]. \tag{29}$$

We can recover q_1 as follows. When $q\in[\underline{q},q_1)$, from (26) it follows

$$\int_{\underline{q}}^{q}\frac{T'(q)-c}{T'(q)}g(q)dq=-\int_{\underline{q}}^{q}\frac{F(\theta)}{\theta f(\theta)}g(q)dq=-\int_{\underline{\theta}}^{\theta}\frac{F(\theta)}{\theta f(\theta)}\frac{f(\theta)}{q'(\theta)}q'(\theta)d\theta=-\int_{\underline{\theta}}^{\theta}\frac{F(\theta)}{\theta}d\theta<0 \tag{30}$$

by substitution, the fact that $f(\theta)=g(q)q'(\theta)$, and that $dq=q'(\theta)d\theta$. Instead, when $q\in[q_1,q_2]$, from (26) and (28) we have

$$\frac{T'(q)-c}{T'(q)}=\frac{\gamma(\theta)-F(\theta)}{\theta f(\theta)}=\frac{F(\theta)-F(\theta_1)}{\theta f(\theta)[F(\theta_2)-F(\theta_1)]}-\frac{F(\theta)}{\theta f(\theta)}=\frac{F(\theta)\left\{1-[F(\theta_2)-F(\theta_1)]\right\}-F(\theta_1)}{\theta f(\theta)[F(\theta_2)-F(\theta_1)]}.$$

From this, again using the rule of integration by substitution, the fact that $g(q)=f(\theta)/q'(\theta)$, and that $dq=q'(\theta)d\theta$, it follows

$$\int_{q_1}^{q}\frac{T'(q)-c}{T'(q)}g(q)dq=\int_{\theta_1}^{\theta}\frac{F(\theta)\left\{1-[F(\theta_2)-F(\theta_1)]\right\}-F(\theta_1)}{\theta[F(\theta_2)-F(\theta_1)]}d\theta>0.$$

Observe that the integral of $[T'(q)-c]g(q)/T'(q)$ over quantities above q_2 is also positive. Then, on the basis of these observations we can conclude

$$q_1=\sup\left\{q\in[\underline{q},\overline{q}]:\int_{\underline{q}}^{q}[T'(q)-c]/T'(q)dG(q)\leq 0\right\}. \tag{31}$$

In practice, (31) is straightforward to compute. The reason is as follows. Since the support of observed purchased quantities is discrete, the observed cumulative distribution function $G(q)$ is a step function, which implies that the integral in (31) simplifies to a summation.

Thus, to recover q_1 and q_2 we follow a three-step procedure:

(1) determine q_1 as in (31);

(2) determine $\widetilde{q}$ as the (interior) quantity greater than q_1 at which $T'(q)$ equals (in practice, is closest to) $T'(\underline{q})$ or $T'(\overline{q})$;
(3) given q_1 and $\widetilde{q}$, use (29) to determine q_2.

With straightforward modifications, the same argument would apply even if $q_1 = \underline{q}$ or $q_2 = \overline{q}$. Once c, q_1, q_2, and $\gamma(\theta)$ are identified, we can identify $\nu'(\cdot)$ and $\theta(\cdot)$ as shown in the following lemma.

Lemma 12. *Suppose the weakly-convex case applies. The first-order condition (15) can be rewritten as follows:*
(a) when $q \in (\underline{q}, q_1]$, we have $\gamma(\theta(q)) = 0$ and

$$\nu'(q) = \frac{T'(q)G(q)^{1-\frac{c}{T'(q)}}}{\theta_1 G(q_1)^{1-\frac{c}{T'(q_1)}}} \exp\left\{c\int_q^{q_1} \frac{\log\left[G(x)\right]T''(x)}{\left[T'(x)\right]^2}dx\right\}; \tag{32}$$

(b) when $q \in [q_1, \widetilde{q})$, we have $\gamma(\theta(q)) < G(q)$ and

$$\nu'(q) = \frac{T'(q)\left\{G(q)\left[G(q_2) - G(q_1) - 1\right] + G(q_1)\right\}^{\frac{G(q_2)-G(q_1)}{G(q_2)-G(q_1)-1}\left[1-\frac{c}{T'(q)}\right]}}{\theta_1\left\{G(q_1)\left[G(q_2) - G(q_1) - 1\right] + G(q_1)\right\}^{\frac{G(q_2)-G(q_1)}{G(q_2)-G(q_1)-1}\left[1-\frac{c}{T'(q_1)}\right]}}$$

$$\cdot \exp\left(-\frac{c\left[G(q_2) - G(q_1)\right]}{G(q_2) - G(q_1) - 1}\int_{q_1}^{q} \frac{\log\left\{G(x)\left[G(q_2) - G(q_1) - 1\right] + G(q_1)\right\}T''(x)}{\left[T'(x)\right]^2}dx\right); \tag{33}$$

(c) when $q \in [\widetilde{q} + \varepsilon, q_2)$, we have $\gamma(\theta(q)) > G(q)$ and

$$\nu'(q) = \frac{T'(q)\left\{[1 - G(q_2) + G(q_1)]G(\widetilde{q} + \varepsilon) - G(q_1)\right\}^{\frac{G(q_2)-G(q_1)}{1-G(q_2)+G(q_1)}\left[1-\frac{c}{T'(\widetilde{q}+\varepsilon)}\right]}}{(\widetilde{\theta} + \varepsilon_\theta)\left\{[1 - G(q_2) + G(q_1)]G(q) - G(q_1)\right\}^{\frac{G(q_2)-G(q_1)}{1-G(q_2)+G(q_1)}\left[1-\frac{c}{T'(q)}\right]}}$$

$$\cdot \exp\left\{\frac{c[G(q_2) - G(q_1)]}{1 - G(q_2) + G(q_1)}\int_{\widetilde{q}+\varepsilon}^{q} \frac{\log\left\{[1 - G(q_2) + G(q_1)]G(x) - G(q_1)\right\}T''(x)}{[T'(x)]^2}\right\} \tag{34}$$

where ε_θ is such that $\widetilde{q} + \varepsilon = q(\widetilde{\theta} + \varepsilon_\theta)$, ε, $\varepsilon_\theta > 0$;
(d) $q \in [q_2, \overline{q})$, we have $\gamma(\theta(q)) = 1$ and

$$\nu'(q) = \frac{T'(q)\left[1 - G(q)\right]^{1-\frac{c}{T'(q)}}}{\theta_2\left[1 - G(q_2)\right]^{1-\frac{c}{T'(q_2)}}} \exp\left\{-c\int_{q_2}^{q} \frac{\log\left[1 - G(x)\right]T''(x)}{\left[T'(x)\right]^2}dx\right\}. \tag{35}$$

In each case, $\theta(q) = T'(q)/\nu'(q)$ for the relevant $\nu'(q)$.

The argument for the identification of $\nu'(\cdot)$ and $\theta(\cdot)$ relies on $\nu'(\cdot)$ and $\theta(\cdot)$ being known or estimable functions of observables. In turn, this fact also implies that $F(\theta)$ and $f(\theta)$ are identified by $G(q)$ and $\theta(q)$. Based on Lemma 12, we can also provide useful restatements of the above conditions and express $\nu'(\cdot)$ and $\theta(\cdot)$ as (known) functions of $T'(q)$, $T''(q)$, $G(q)$, c, $\underline{\theta}$, and $\underline{q}$. See Appendix B for proof and details.

Observe that $\nu'(\cdot)$ and $\theta(\cdot)$ are identified up to θ_1 and θ_2. We now show how we only need to normalize $\underline{\theta}$ in order to identify (and estimate) θ_1 and θ_2. To see this, start by normalizing $\underline{\theta}$ to one. In practice, the empirical distribution function $G(q)$ is a step function with steps at $q_1, \ldots, q_N$ in $(\underline{q}, q_1]$. On each of these intervals $G(\cdot)$ is constant whereas the integral of $T''(\cdot)/[T'(\cdot)]^2$ is $-1/T'(\cdot)$. By the above argument, the integral in (32) can be rewritten as the finite sum of integrals from q_i to q_{i+1} where $i = 1, \ldots, i_1 - 1$ and i_1 is the index of the observed quantity q_{i_1} with associated type θ_{i_1}. Formally, let

$$q_{i_1} = \max\{q_i \in \{q_1, \ldots, q_N\} : \sum_{j=1}^{i} [T'(q_j) - c]/T'(q_j)\,\Pr(q_j) \leq 0\}$$

where $\Pr(q_j) = G(q_j) - G(q_{j-1})$. Thus, $\nu'(q)$ in (32) can be rewritten as

$$\nu'(q) = \frac{T'(q)G(q_i)^{1-\frac{c}{T'(q)}}}{\theta_{i_1}G(q_{i_1})^{1-\frac{c}{T'(q_{i_1})}}} \exp\left\{c\sum_{j=i+1}^{i_1-1} \log\left[G(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right.$$

$$\left. + c\log\left[G(q_i)\right]\left[\frac{1}{T'(q)} - \frac{1}{T'(q_{i+1})}\right]\right\}$$

for $q \in [q_i, q_{i+1})$, $i = 1, \ldots, i_1 - 1$. In particular, note that $G(q_1) = \Pr(q \leq q_1) = \Pr(q = q_1) > 0$, so

$$\nu'(q_1) = \frac{T'(q_1)G(q_1)^{1-\frac{c}{T'(q_1)}}}{\theta_{i_1}G(q_{i_1})^{1-\frac{c}{T'(q_{i_1})}}} \exp\left\{c\sum_{j=1}^{i_1-1} \log\left[G(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}.$$

From the fact that $\theta_1\nu'(q_1) = T'(q_1)$ and that $\underline{\theta} = \theta_1 = 1$, we can express θ_{i_1} as

$$\theta_{i_1} = \frac{G(q_1)^{1-\frac{c}{T'(q_1)}}}{G(q_{i_1})^{1-\frac{c}{T'(q_{i_1})}}} \exp\left\{c\sum_{j=1}^{i_1-1} \log\left[G(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}.$$

Consider now the interval $[q_1, q_{\tilde{i}})$. Since $G(\cdot)$ is in practice a discrete distribution, $q_{\tilde{i}}$ is defined as $q_{\tilde{i}} = \max\{q \in \{q_1, \ldots, q_N\} : \gamma(\theta(q_i)) \leq G(q)\}$. In general, no point in the support of $G(\cdot)$ satisfies $\gamma = G(q)$. So, $\nu'(q)$ can be computed exactly over its entire support, that is, over the intervals $[q_i, q_{i+1})$, $i \in \{i_1, \ldots, \widetilde{i} - 1\}$, and over the intervals $[q_i, q_{i+1})$, $i \in \{\widetilde{i}, \ldots, N-1\}$. In other words, ε and ε_θ in (34) can be taken to be zero. (In the knife-edged case in which there exists a point in the support of $G(\cdot)$ such that $\gamma = G(q)$, note that $\nu'(q)$ can still be computed exactly but only over the intervals $[q_i, q_{i+1})$, $i \in \{i_1, \ldots, \widetilde{i}-1\}$, and $[q_i, q_{i+1})$, $i \in \{\widetilde{i}+1, \ldots, N-1\}$.) Let q_{i_2} be the empirical counterpart of q_2 (see Appendix C for details). From Appendix B we know that if $q \in [q_1, \widetilde{q})$ then $\gamma(\theta(q)) < G(q)$ and $\nu'(q)$ can be written as

$$\nu'(q) = \frac{T'(q)\left\{G(q_i)\left[G(q_{i_2}) - G(q_{i_1}) - 1\right] + G(q_{i_1})\right\}^{\frac{G(q_{i_2})-G(q_{i_1})}{G(q_{i_2})-G(q_{i_1})-1}\left[1-\frac{c}{T'(q)}\right]}}{\theta_{i_1}\left\{G(q_{i_1})\left[G(q_{i_2}) - G(q_{i_1}) - 1\right] + G(q_{i_1})\right\}^{\frac{G(q_{i_2})-G(q_{i_1})}{G(q_{i_2})-G(q_{i_1})-1}\left[1-\frac{c}{T'(q_{i_1})}\right]}}$$

$$\cdot \exp\left(-\frac{c\left[G(c) - G(q_{i_1})\right]}{G(q_{i_2}) - G(q_{i_1}) - 1}\sum\nolimits_{j=i_1}^{i-1} \log\left\{G(q_j)\left[G(q_{i_2}) - G(q_{i_1}) - 1\right] + G(q_{i_1})\right\}\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right.$$
$$\left. -\frac{c\left[G(q_{i_2}) - G(q_{i_1})\right]}{G(q_{i_2}) - G(q_{i_1}) - 1}\log\left\{G(q_i)\left[G(q_{i_2}) - G(q_{i_1}) - 1\right] + G(q_{i_1})\right\}\left[\frac{1}{T'(q_i)} - \frac{1}{T'(q)}\right]\right)$$

for $q \in [q_i, q_{i+1})$ where $i = i_1, \ldots, \widetilde{i} - 1$. Note that when $q \uparrow q_{\widetilde{i}}$ with $q \in [q_{\widetilde{i}-1}, q_{\widetilde{i}})$ we have

$$\lim_{q \uparrow q_{\widetilde{i}}}\{G(q)\left[G(q_{i_2}) - G(q_{i_1}) - 1\right] + G(q_{i_1})\} = G(q_{\widetilde{i}-1})\left[G(q_{i_2}) - G(q_{i_1}) - 1\right] + G(q_{i_1})$$

since $G(\cdot)$ is continuous from the right, whereas $\lim_{q\uparrow q_{\widetilde{i}}}[1 - c/T'(q)] \approx 0$. (This in an approximation: the limit is exactly zero if there exists a point in the support of $G(\cdot)$ such that $\gamma = G(q)$.) So,

$$\lim_{q\uparrow q_{\widetilde{i}}} \nu'(q) \approx \frac{T'(\widetilde{q}_i)}{\theta_{i_1}\left\{G(q_{i_1})\left[G(q_{i_2}) - G(q_{i_1}) - 1\right] + G(q_{i_1})\right\}^{\frac{G(q_{i_2})-G(q_{i_1})}{G(q_{i_2})-G(q_{i_1})-1}\left[1-\frac{c}{T'(q_{i_1})}\right]}}$$

$$\cdot \exp\left(-\frac{c\left[G(q_{i_2}) - G(q_{i_1})\right]}{G(q_{i_2}) - G(q_{i_1}) - 1}\sum\nolimits_{j=i_1}^{\widetilde{i}-1} \log\left\{G(q_j)\left[G(q_{i_2}) - G(q_{i_1}) - 1\right] + G(q_{i_1})\right\}\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right).$$

From this, the fact that $\widetilde{\theta}_i\nu'(\widetilde{q}_i) = T'(\widetilde{q}_i)$, and the expression for θ_{i_1}, we can also identify $\widetilde{\theta}_i$.

Now, from Appendix B when $q \in [q_{\widetilde{i}}, q_{i_2})$ we know that $\gamma(\theta(q)) > G(q)$ and that $\nu'(q)$ in (34) can alternatively be expressed as

$$\nu'(q) = \frac{T'(q)\left\{[1 - G(q_{i_2}) + G(q_{i_1})]G(\widetilde{q}_i + \varepsilon) - G(q_{i_1})\right\}^{\frac{G(q_{i_2})-G(q_{i_1})}{1-G(q_{i_2})+G(q_{i_1})}\left[1-\frac{c}{T'(\widetilde{q}_i+\varepsilon)}\right]}}{(\widetilde{\theta}_i + \varepsilon_\theta)\left\{[1 - G(q_{i_2}) + G(q_{i_1})]G(q_i) - G(q_{i_1})\right\}^{\frac{G(q_{i_2})-G(q_{i_1})}{1-G(q_{i_2})+G(q_{i_1})}\left[1-\frac{c}{T'(q)}\right]}}$$

$$\cdot \exp\left\{\frac{c[G(q_{i_2}) - G(q_{i_1})]}{1 - G(q_{i_2}) + G(q_{i_1})}\sum\nolimits_{j=\widetilde{i}}^{i-1}\log\left\{[1 - G(q_{i_2}) + G(q_{i_1})]G(q_j) - G(q_{i_1})\right\}\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right.$$

$$+\frac{c[G(q_{i_2})-G(q_{i_1})]}{1-G(q_{i_2})+G(q_{i_1})}\log\{[1-G(q_{i_2})+G(q_{i_1})]G(q_i)-G(q_{i_1})\}\left[\frac{1}{T'(q_i)}-\frac{1}{T'(q)}\right]\Bigg\} \tag{36}$$

for $q \in [q_i, q_{i+1})$ where $i = \widetilde{i}, \ldots, i_2 - 1$ and i_2 is the index of the quantity q_{i_2} corresponding to q_2. Given $\theta_{i_2} = T'(q_{i_2})/\nu'(q_{i_2})$, by the above argument we can compute

$$\nu'(q_{i_2}) = \frac{T'(q_{i_2})\{[1-G(q_{i_2})+G(q_{i_1})]G(\widetilde{q}_i)-G(q_{i_1})\}^{\frac{G(q_{i_2})-G(q_{i_1})}{1-G(q_{i_2})+G(q_{i_1})}\left[1-\frac{c}{T'(\widetilde{q}_i+\varepsilon)}\right]}}{(\widetilde{\theta}_i+\varepsilon_\theta)\{[1-G(q_{i_2})+G(q_{i_1})]G(q_{i_2})-G(q_{i_1})\}^{\frac{G(q_{i_2})-G(q_{i_1})}{1-G(q_{i_2})+G(q_{i_1})}\left[1-\frac{c}{T'(q_{i_2})}\right]}}$$

$$\cdot \exp\left\{\frac{c[G(q_{i_2})-G(q_{i_1})]}{1-G(q_{i_2})+G(q_{i_1})}\sum_{j=\widetilde{i}}^{i_2-1}\log\{[1-G(q_{i_2})+G(q_{i_1})]G(q_j)-G(q_{i_1})\}\left[\frac{1}{T'(q_j)}-\frac{1}{T'(q_{j+1})}\right]\right\}$$

and obtain θ_{i_2} from $\theta_{i_2} = T'(q_{i_2})/\nu'(q_{i_2})$. Lastly, when $q \in [q_{i_2}, q_N)$ we can show that $\nu'(q)$ can be rewritten as

$$\nu'(q) = \frac{T'(q)\left[1-G(q_i)\right]^{1-\frac{c}{T'(q)}}}{\theta_{i_2}\left[1-G(q_{i_2})\right]^{1-\frac{c}{T'(q_{i_2})}}}\exp\Bigg\{-c\sum_{j=i_2}^{i-1}\log\left[1-G(q_j)\right]\left[\frac{1}{T'(q_j)}-\frac{1}{T'(q_{j+1})}\right]$$

$$-c\log\left[1-G(q_i)\right]\left[\frac{1}{T'(q_i)}-\frac{1}{T'(q)}\right]\Bigg\}$$

for $q \in [q_i, q_{i+1})$ where $i = i_2, \ldots, N-1$. Note that when $q \uparrow q_N$ with $q \in [q_{N-1}, q_N)$, we have $G(q) = G(q_{N-1}) = \Pr(q \le q_{N-1}) \in (0,1)$ whereas $\lim_{q\uparrow q_N}[1 - c/T'(q)] = 0$, so

$$\lim_{q\uparrow q_N}\nu'(q) = \frac{T'(q_N)\exp\left\{-c\sum_{j=i_2}^{N-1}\log\left[1-G(q_j)\right]\left[\frac{1}{T'(q_j)}-\frac{1}{T'(q_{j+1})}\right]\right\}}{\theta_{i_2}\left[1-G(q_{i_2})\right]^{1-\frac{c}{T'(q_{i_2})}}}$$

and from $\theta_N \nu'(q_N) = T'(q_N)$, θ_N can be obtained in a straightforward manner.

8.2.2 Highly-Convex Case

Recall that in the highly-convex case, (IR) binds for one or both of the extreme types $\underline{\theta}$ and $\overline{\theta}$, and $\gamma(\theta)$ is constant for all types. Consider first the case in which $\gamma(\theta) = \gamma = 1$, as in the standard model. Then, from (26) it follows

$$T'(q) \ge c = T'(\overline{q}) \tag{37}$$

with equality at $\overline{q}$, since $F(\theta) \leq 1$, $f(\theta) > 0$ by assumption, and $\theta > 0$ by (IA5). Consider now the case in which $\gamma(\theta) = \gamma = 0$. From (26) it follows

$$T'(q) \leq T'(\underline{q}) = c \tag{38}$$

with equality at $\underline{q}$, since $F(\theta) \geq 0$, $f(\theta) > 0$ by assumption, and $\theta > 0$ by (IA5). Consider now the general case in which $\gamma \in (0, 1)$. By the same argument, (26) implies that $T'(\underline{q}) - c > 0 > T'(\overline{q}) - c$, which yields $T'(\underline{q}) > c > T'(\overline{q})$. Rewrite (26) as

$$T'(q) = c + \frac{1}{\theta q'(\theta)} \frac{[\widehat{\gamma}(\theta(q)) - G(q)]T'(q)}{g(q)}$$

so we can specify an additive nonparametric regression model in which the partial regression function $f_R(\cdot)$ is estimated from the data

$$T'(q) = c + f_R\left(\frac{[\widehat{\gamma}(\theta(q)) - G(q)]T'(q)}{g(q)}\right). \tag{39}$$

Hence, if we regress $T'(q)$ on a constant and $T'(q)/g(q)$, the value of the estimated constant provides an estimate of c. Again, q can be treated as exogenous given that all relevant functions are deterministic. Recall from (27) we can also identify (and estimate) γ (DICUSS BOUNDARY CONDITIONS).

Lastly, recall that $\widetilde{q}$ denotes the quantity in $[\underline{q}, \overline{q}]$ such that γ equals $G(q)$. Formally, the quantity $\widetilde{q}$ is defined as $\widetilde{q} = \inf\{q \in [\underline{q}, \overline{q}] : \gamma \leq G(q)\}$. Once γ is determined, $\widetilde{q}$ can be easily determined from $G(q)$. We summarize these observations in the following lemma.

Lemma 13. *Suppose the highly-convex case applies. The estimated constant of a regression of $G(q)$ on $g(q)/T'(q)$ provides an estimate of γ whereas the estimated constant of a regression of $T'(q)$ on $[\widehat{\gamma}(\theta(q)) - G(q)]T'(q)/g(q)$ provides an estimate of c. The first-order condition (15) is equivalent to*

$$\nu'(q) = \frac{T'(q)}{\underline{\theta}} \exp\left\{-\int_{\underline{q}}^{q} \frac{g(x)}{\gamma - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx\right\} \tag{40}$$

when $q \in [\underline{q}, \widetilde{q})$, $\widetilde{q} = \inf\{q \in [\underline{q}, \overline{q}] : \gamma \leq G(q)\}$, and to

$$\nu'(q) = \frac{T'(q)}{(\widetilde{\theta} + \varepsilon_\theta)} \exp\left\{\int_{\widetilde{q}+\varepsilon}^{q} \frac{g(x)}{G(x) - \gamma}\left[1 - \frac{c}{T'(x)}\right] dx\right\} \tag{41}$$

when $q \in [\widetilde{q} + \varepsilon, \overline{q})$, where $\widetilde{q} + \varepsilon = q(\widetilde{\theta} + \varepsilon_\theta)$ with ε, $\varepsilon_\theta > 0$. In each case, $\theta(q) = T'(q)/\nu'(q)$ for *the relevant $\nu'(q)$.*

As in the weakly-convex case, based on this lemma we can provide useful restatements of the above conditions so that $\nu'(\cdot)$ and $\theta(\cdot)$ can be expressed as known functions of $T'(q)$, $T''(q)$, $G(q)$, c, $\underline{\theta}$, $\underline{q}$, $\widetilde{\theta}$, and $\widetilde{q}$. See Appendix B for details.

Since $G(\cdot)$ is in practice a discrete distribution, $\widetilde{q}_i$ is defined as $\widetilde{q}_i = \min\{q \in \{q_1, \ldots, q_N\} : \gamma \leq G(q_i)\}$. Then, in general, no point in the support of $G(\cdot)$ satisfies $\gamma = G(q_i)$. So, $\nu'(q)$ can be computed exactly over its entire support, that is, over the intervals $[q_i, q_{i+1})$, $i \in \{i_1, \ldots, \widetilde{i} - 1\}$, according to (40) and over the intervals $[q_i, q_{i+1})$, $i \in \{\widetilde{i}, \ldots, N-1\}$, according to (41). In other words, ε and ε_θ in (41) can be taken to be zero. In the knife-edged case in which there exists a point in the support of $G(\cdot)$ such that $\gamma = G(q)$, observe that $\nu'(q)$ can still be computed exactly but only over the intervals $[q_i, q_{i+1})$, $i \in \{i_1, \ldots, \widetilde{i} - 1\}$, according to (40) and over the intervals $[q_i, q_{i+1})$, $i \in \{\widetilde{i} + 1, \ldots, N-1\}$, according to (41).

As before, we manipulate the conditions in Lemma 13 to obtain convenient estimators of the model's primitives. We proceed by first normalizing $\underline{\theta}$ to one. Recall that we order the observed quantities purchased in a market from smallest, $\underline{q} = q_1$, to largest, $\overline{q} = q_N$. For any $\gamma > 0$, when $q \in [q_1, \widetilde{q}_i)$, we know that $\gamma > G(q)$, so we can express $\nu'(q)$ from (40) as

$$\nu'(q) = \frac{T'(q)\left[\gamma - G(q_i)\right]^{1-\frac{c}{T'(q)}}}{\gamma^{1-\frac{c}{T'(q_1)}}} \exp\left\{-c\sum_{j=1}^{i-1} \log\left[\gamma - G(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right.$$
$$\left. -c\log\left[\gamma - G(q_i)\right]\left[\frac{1}{T'(q_i)} - \frac{1}{T'(q)}\right]\right\}$$

for $q \in [q_i, q_{i+1})$, $i = 1, \ldots, \widetilde{i} - 1$. Since

$$\lim_{q\uparrow\widetilde{q}_i}\left[\gamma - G(q)\right]^{1-\frac{c}{T'(q)}} = \lim_{q\uparrow\widetilde{q}_i}\left[\gamma - G(\widetilde{q}_{i-1})\right]^{1-\frac{c}{T'(q)}} \approx 1$$

it follows

$$\lim_{q\uparrow\widetilde{q}_i}\nu'(q) \approx \frac{T'(q_{\widetilde{i}})}{\gamma^{1-\frac{c}{T'(q_1)}}} \exp\left\{-c\sum_{j=1}^{\widetilde{i}-1} \log\left[\gamma - G(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}$$

and using the fact that $\widetilde{\theta}_i\nu'(\widetilde{q}_i) = T'(\widetilde{q}_i)$, we obtain

$$\widetilde{\theta}_i \approx \gamma^{1-\frac{c}{T'(q_1)}} \exp\left\{c\sum_{j=1}^{\widetilde{i}-1} \log\left[\gamma - G(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}.$$

When $q \in [\widetilde{q}_i, \overline{q}]$ (ignoring for simplicity the knife-edged case in which $G(\widetilde{q}_i) = \gamma$) but $\gamma < 1$, we know that $\gamma < G(q)$ and we can express $\nu'(q)$ from (41) also as

$$\nu'(q) = \frac{T'(q)[G(q_i) - \gamma]^{1-\frac{c}{T'(q)}}}{\widetilde{\theta}_i[G(\widetilde{q}_i) - \gamma]^{1-\frac{c}{T'(\widetilde{q}_i)}}} \exp\left\{-c\sum_{j=\widetilde{i}}^{i-1} \log\left[G(q_j) - \gamma\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right.$$
$$\left. -c\log\left[G(q_i) - \gamma\right]\left[\frac{1}{T'(q_i)} - \frac{1}{T'(q)}\right]\right\}$$

with $q \in [q_i, q_{i+1})$, $i = \widetilde{i}, \ldots, N-1$ and

$$\nu'(q_N) = \frac{T'(q_N)[G(q_N) - \gamma]^{1-\frac{c}{T'(q_N)}}}{\widetilde{\theta}_i[G(\widetilde{q}_i) - \gamma]^{1-\frac{c}{T'(\widetilde{q}_i)}}} \exp\left\{-c\sum_{j=\widetilde{i}}^{N-1} \log\left[G(q_j) - \gamma\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}.$$

From this, $\overline{\theta} = T'(q_N)/\nu'(q_N)$ immediately follows. Now, if $\gamma = 1$, then for $q \uparrow q_N$ with $q \in [q_{N-1}, q_N)$ we have $G(q) = G(q_{N-1}) = \Pr(q \leq q_{N-1}) \in (0,1)$ whereas $\lim_{q\uparrow q_N}[1 - c/T'(q)] = 0$. Therefore,

$$\lim_{q\uparrow q_N} \nu'(q) = T'(q_N)\exp\left\{-c\sum_{j=1}^{N-1} \log\left[1 - G(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}$$

and using the fact that $\overline{\theta}\nu'(q_N) = T'(q_N)$, we obtain

$$\overline{\theta} = \exp\left\{c\sum_{j=1}^{N-1} \log\left[1 - G(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}.$$

Lastly, note that the case in which $\gamma = 0$ is just a special case of this with $q_{\widetilde{i}} = q_1$. From these derivations, it follows that $f(\theta)$ is also straightforwardly identified.

8.3 Nonparametric Estimation

The parameters and functions of the model to be recovered are: $\nu(\cdot)$, $F(\cdot)$, c, and $\overline{u}(\cdot)$. Estimation involves three steps and is based on GPV and PV. In the first step, we estimate the distribution of types for which (IR) binds $\gamma(\theta)$, the (constant) marginal cost c, and the distribution of quantity purchases $G(q)$. Based on the recovered $\gamma(\theta)$, we can determine $\overline{q}(\theta)$ and, by comparison with $q(\theta)$, whether or not budget constraints are likely to matter. In the second step, we estimate the base marginal utility functions $\nu'(q)$ and $\overline{u}'(\cdot)$. Different estimators can be obtained depending on the type of optimal contract identified in the previous step. In the third step, we estimate the distribution of consumer types $F(\theta)$ based on the sample of pseudo-types derived in the previous step from the structural relationship $\theta(q) = T'(q)/\nu'(q)$ (local incentive compatibility). We now turn to discuss each step in more detail. We denote an estimated quantity x by $\widehat{x}$.

Price Schedule. In our data we have direct information on unit prices and quantity purchased by each consumer (household) in a market (village). Then, the price schedule can be simply obtained by fitting observed prices on quantities, for instance by least squares,

$$\widehat{\log(T)} = \hat{t}_0 + \hat{t}_1 \log(q) + \hat{t}_2(\log(q))^2.$$

The reason the quantity q can be treated as exogenous is that having information on the quantity purchased by each consumer is equivalent to having information on the price schedule of the seller (see PV). Also, according to the model, $T(q)$ is a deterministic function. If the fit is good, then the measurement error that fitting may cause can be considered minimal and, thus, ignored.

Contract Type. By Lemma 9, we can determine whether the price schedule observed in each market, as constructed above, corresponds to the weakly-convex case or the highly-convex case, a special case being the standard case with $\gamma(\theta) = 1$ for all θ. Depending on the case that proves relevant, we estimate $\gamma(\theta)$ as detailed above.

Quantity Distribution. Denote by N the number of consumers purchasing a positive quantity and by q_i the quantity purchased by consumer i. Note that $G(q)$ can be estimated as an empirical distribution using a counting process as in PV, where for $q \in [\underline{q}, \overline{q}]$,

$$\widehat{G}(q) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(q_i \leq q)$$

and $\mathbf{1}(\cdot)$ is an indicator function. Since, as argued, $\widehat{G}(q)$ is a step function, $\widehat{g}(q)$ can also be easily estimated.[9] Note that, as mentioned at the end of Lemmas 12 and 13, alternative estimators of $\nu'(q_i)$ and $\theta(q_i)$ can be derived that do *not* require estimation of $\widehat{g}(q)$. See Appendix B for details.

Base Marginal Utility Function. The estimation of $\nu'(q_i)$ and $\theta(q_i)$ in the weakly-convex case and in the highly-convex case follows immediately from Lemmas 12 and 13. Given our homogeneity assumption, at this step we can also recover $\overline{u}'(\cdot)$.

Type Distribution. Following GPV and PV, given an estimate for $\widehat{\nu}'(q_i)$, a sample of pseudo-

[9] Alternative nonparametric smoothed kernel density/estimator methods are also possible.

types can be constructed from $\widehat{\theta}_i = \widehat{\theta}(q_i) = T'(q_i)/\widehat{\nu}'(q_i)$. Then, the density $f(\theta)$ can be estimated nonparametrically as

$$\widehat{f}(\theta) = \frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{\theta - \widehat{\theta}_i}{h}\right)$$

for a suitable choice of (symmetric) kernel function $K(\cdot)$ and bandwidth h. Additional details are provided in Appendix C.

9 Conclusion

We have proposed a model of nonlinear pricing in which consumers have private information about their taste for a good, their consumption possibilities, and their budgets. In this case the properties of an optimal allocation under adverse selection can be fundamentally different from those of an optimal allocation according to the standard nonlinear pricing model, in which consumers are assumed unconstrained in their purchasing choices. In particular, in our environment quantity discounts for *large* volumes can be associated with overprovision of quantity at *low* volumes. Depending on the shape of consumers' reservation utility profile, overproduction can occur for high types or for low types or for all types of consumers. In these latter two cases, nonlinear pricing can positively affect the purchasers of the smallest quantities, typically the poorest consumers.

We have shown that a nonlinear pricing problem in which consumers are budget-constrained is an instance of one in which consumers are not budget-constrained but their outside option to trading with a particular seller depends on their private information. We have proved that this more general model in which reservation utility is type-dependent is nonparametrically identified. We have also derived nonparametric estimators of the model primitives that can be readily implemented using publicly available data from conditional cash transfer programs, common in several developing countries. Hence, this framework provides a useful structure to assess the impact of asymmetric information, outside options, and budget constraints on consumer and producer surplus.

A Appendix A

Proof of Lemma 1: (a) The argument is an application of the proofs of Lemma 1 and Lemma 2 in Jullien (2000). We first prove the 'only if' part of the claim. Under the assumptions of the lemma, there exists a tariff $\overline{c}(q)$ such that, when faced with it, the consumer chooses to purchase $\overline{q}(\theta)$ and obtains the utility $\overline{u}(\theta)$. Let $t(q)$ be a tariff that implements the allocation $\{u(\theta), q(\theta), x(\theta)\}$. If the seller offers the consumer the tariff $\inf\{t(q), \overline{c}(q)\}$, then it is optimal for the consumer to participate, choose the quantity $x(\theta)q(\theta) + [1 - x(\theta)]\overline{q}(\theta)$, and obtain the utility $\max\{u(\theta), \overline{u}(\theta)\} = u(\theta)$. The 'if part' of the claim is immediate, as it amounts to excluding some types receiving their reservation utility, which cannot affect incentives. (b) Consider an implementable allocation $\{u(\theta), q(\theta), x(\theta)\}$ and the full participation allocation $\{u(\theta), x(\theta)q(\theta) + (1 - x(\theta))\overline{q}(\theta)\}$. The welfare levels coincide in the two allocations for a type θ who participates. If $x(\theta) = 0$, then the seller's profit from this type under the full participation allocation is $s(\theta, \overline{q}(\theta)) \geq \overline{u}(\theta)$. So, the full participation allocation weakly dominates the allocation with exclusion. Therefore, any optimal allocation must include all types. □

Weakly-Convex and Highly-Convex Case: Applying the implicit function theorem to (13) on Θ, it follows

$$l_\theta(\overline{\gamma}(\theta), \theta) = \frac{\partial l(\overline{\gamma}(\theta), \theta)}{\partial \theta} = -\frac{\partial \sigma_q(\overline{\gamma}(\theta), \theta, q)/\partial \theta}{\partial \sigma_q(\overline{\gamma}(\theta), \theta, q)/\partial q} =$$

$$-\frac{s_{q\theta}(\theta, l(\overline{\gamma}(\theta), \theta)) + \left\{\frac{f^2(\theta) - \overline{\gamma}'(\theta)f(\theta) - [F(\theta) - \overline{\gamma}(\theta)]f'(\theta)}{f^2(\theta)}\right\} v_{\theta q}(\theta, l(\overline{\gamma}(\theta), \theta)) + \frac{F(\theta) - \overline{\gamma}(\theta)}{f(\theta)} v_{\theta\theta q}(\theta, l(\overline{\gamma}(\theta), \theta))}{s_{qq}(\theta, l(\overline{\gamma}(\theta), \theta)) + \frac{F(\theta) - \overline{\gamma}(\theta)}{f(\theta)} v_{\theta qq}(\theta, l(\overline{\gamma}(\theta), \theta))}.$$

Substituting $\overline{\gamma}(\theta) = F(\theta) + f(\theta)\frac{s_q(\theta, \overline{q}(\theta))}{v_{\theta q}(\theta, \overline{q}(\theta))}$ from (11), we obtain $l_\theta(\overline{\gamma}(\theta), \theta)$ equals

$$-\frac{s_{q\theta}(\theta, l(\overline{\gamma}(\theta), \theta)) + \left[1 - \frac{\overline{\gamma}'(\theta)}{f(\theta)} + \frac{f'(\theta)}{f(\theta)} \cdot \frac{s_q(\theta, \overline{q}(\theta))}{v_{\theta q}(\theta, \overline{q}(\theta))}\right] v_{\theta q}(\theta, l(\overline{\gamma}(\theta), \theta)) - \frac{s_q(\theta, \overline{q}(\theta))}{v_{\theta q}(\theta, \overline{q}(\theta))} v_{\theta\theta q}(\theta, l(\overline{\gamma}(\theta), \theta))}{s_{qq}(\theta, l(\overline{\gamma}(\theta), \theta)) - \frac{s_q(\theta, \overline{q}(\theta))}{v_{\theta q}(\theta, \overline{q}(\theta))} v_{\theta qq}(\theta, l(\overline{\gamma}(\theta), \theta))}.$$

With $s_{q\theta}(\theta, q) = v_{\theta q}(\theta, q)$ and $s_{qq}(\theta, q) = v_{qq}(\theta, q) - c''(q)$, we can rewrite $l_\theta(\overline{\gamma}(\theta), \theta)$ as

$$l_\theta(\overline{\gamma}(\theta), \theta) = -\frac{\left[2 - \frac{\overline{\gamma}'(\theta)}{f(\theta)} + \frac{f'(\theta)}{f(\theta)} \cdot \frac{s_q(\theta, \overline{q}(\theta))}{v_{\theta q}(\theta, \overline{q}(\theta))}\right] v_{\theta q}(\theta, l(\overline{\gamma}(\theta), \theta)) - \frac{s_q(\theta, \overline{q}(\theta))}{v_{\theta q}(\theta, \overline{q}(\theta))} v_{\theta\theta q}(\theta, l(\overline{\gamma}(\theta), \theta))}{v_{qq}(\theta, l(\overline{\gamma}(\theta), \theta)) - c''(l(\overline{\gamma}(\theta), \theta)) - \frac{s_q(\theta, \overline{q}(\theta))}{v_{\theta q}(\theta, \overline{q}(\theta))} v_{\theta qq}(\theta, l(\overline{\gamma}(\theta), \theta))}$$

$$= -\frac{\left[\left(2 - \frac{\overline{\gamma}'(\theta)}{f(\theta)}\right) v_{\theta q}(\theta, \overline{q}(\theta)) + \frac{f'(\theta)s_q(\theta, \overline{q}(\theta))}{f(\theta)}\right] v_{\theta q}(\theta, l(\overline{\gamma}(\theta), \theta)) - s_q(\theta, \overline{q}(\theta)) v_{\theta\theta q}(\theta, l(\overline{\gamma}(\theta), \theta))}{v_{\theta q}(\theta, \overline{q}(\theta)) \left[v_{qq}(\theta, l(\overline{\gamma}(\theta), \theta)) - c''(l(\overline{\gamma}(\theta), \theta))\right] - s_q(\theta, \overline{q}(\theta)) v_{\theta qq}(\theta, l(\overline{\gamma}(\theta), \theta))}.$$

Finally, substituting $\overline{q}(\theta)$ for $l(\overline{\gamma}(\theta), \theta)$, we obtain $l_\theta(\overline{\gamma}(\theta), \theta)$ equals

$$-\frac{\left[\left(2 - \frac{\overline{\gamma}'(\theta)}{f(\theta)}\right) v_{\theta q}(\theta, \overline{q}(\theta)) + f'(\theta)s_q(\theta, \overline{q}(\theta))/f(\theta)\right] v_{\theta q}(\theta, \overline{q}(\theta)) - s_q(\theta, \overline{q}(\theta)) v_{\theta\theta q}(\theta, \overline{q}(\theta))}{v_{\theta q}(\theta, \overline{q}(\theta)) \left[v_{qq}(\theta, \overline{q}(\theta)) - c''(\overline{q}(\theta))\right] - s_q(\theta, \overline{q}(\theta)) v_{\theta qq}(\theta, \overline{q}(\theta))}.$$

Instead, by applying again the implicit function theorem to $\overline{s}_q(\theta, q) = v_q(\theta, q) - \overline{c}'(q) = 0$, we obtain

$$\frac{\partial \overline{q}(\theta)}{\partial \theta} = -\frac{\partial \overline{s}_q(\theta, q)/\partial \theta}{\partial \overline{s}_q(\theta, q)/\partial q} = -\frac{v_{\theta q}(\theta, q)}{v_{qq}(\theta, q) - \overline{c}''(q)} = -\frac{v_{\theta q}(\theta, \overline{q}(\theta))}{v_{qq}(\theta, \overline{q}(\theta)) - \overline{c}''(\overline{q}(\theta))}.$$

Hence, the weakly-convex case applies if

$$l_\theta(\overline{\gamma}(\theta),\theta) = \frac{\partial l(\overline{\gamma}(\theta),\theta)}{\partial\theta} \geq \frac{\partial \overline{q}(\theta)}{\partial\theta} = \overline{q}'(\theta)$$

which is satisfied if, and only if,

$$-\frac{\left(2-\frac{\overline{\gamma}'(\theta)}{f(\theta)}\right) v_{\theta q}(\theta,\overline{q}(\theta)) + f'(\theta)s_q(\theta,\overline{q}(\theta))/f(\theta) - s_q(\theta,\overline{q}(\theta))v_{\theta\theta q}(\theta,\overline{q}(\theta))/v_{\theta q}(\theta,\overline{q}(\theta))}{v_{\theta q}(\theta,\overline{q}(\theta))\left[v_{qq}(\theta,\overline{q}(\theta)) - c''(\overline{q}(\theta))\right] - s_q(\theta,\overline{q}(\theta))v_{\theta qq}(\theta,\overline{q}(\theta))} \geq$$

$$-\frac{1}{v_{qq}(\theta,\overline{q}(\theta)) - \overline{c}''(\overline{q}(\theta))}. \tag{42}$$

Consider our leading example, with $v(\theta,q)=\theta\log(q)$, $I(\theta)=I+i\theta$. By (7),

$$\overline{c}(q) = I + i\underline{\theta}\exp\left\{\frac{\underline{\theta}\nu(q) - I - i\underline{\theta} - \overline{u}}{i\underline{\theta}}\right\} = I + i\underline{\theta}\exp\left\{-\frac{I+i\underline{\theta}+\overline{u}}{i\underline{\theta}}\right\} q^{\frac{1}{i}}$$

so

$$\overline{c}'(q) = \underline{\theta}\exp\left\{-\frac{I+i\underline{\theta}+\overline{u}}{i\underline{\theta}}\right\} q^{\frac{1-i}{i}}$$

and

$$\overline{c}''(q) = \left(\frac{1-i}{i}\right)\underline{\theta}\exp\left\{-\frac{I+i\underline{\theta}+\overline{u}}{i\underline{\theta}}\right\} q^{\frac{1-2i}{i}}.$$

Also, by (8)

$$\overline{q}(\theta) = (\nu)^{-1}\left(i\log\left(\frac{\theta}{\underline{\theta}}\right) + \frac{I+i\underline{\theta}+\overline{u}}{\underline{\theta}}\right) = (\nu)^{-1}\left(\log\left[\left(\frac{\theta}{\underline{\theta}}\right)^i\right] + \log\left[\exp\left(\frac{I+i\underline{\theta}+\overline{u}}{\underline{\theta}}\right)\right]\right)$$

it follows

$$\overline{q}(\theta) = \exp\left(\frac{I+i\underline{\theta}+\overline{u}}{\underline{\theta}}\right)\left(\frac{\theta}{\underline{\theta}}\right)^i$$

and

$$\overline{c}''(\overline{q}(\theta)) = \left(\frac{1-i}{i}\right)\underline{\theta}\exp\left[-\frac{2(I+i\underline{\theta}+\overline{u})}{\underline{\theta}}\right]\left(\frac{\theta}{\underline{\theta}}\right)^{1-2i}.$$

Assume that θ is uniformly distributed on $[\underline{\theta},\overline{\theta}]$, so $f(\theta)=1/(\overline{\theta}-\underline{\theta})$ and $F(\theta)=(\theta-\underline{\theta})/(\overline{\theta}-\underline{\theta})$, and that the seller's marginal cost is constant. Note that $v_\theta(\theta,q)=\log(q)$, $v_{\theta\theta}(\theta,q)=0$, $v_{\theta\theta q}(\theta,q)=0$, $v_{\theta q}(\theta,q)=1/q$, $v_{\theta qq}(\theta,q)=-1/q^2$, $v_q(\theta,q)=\theta/q$, and $v_{qq}(\theta,q)=-\theta/q^2=\theta v_{\theta qq}(\theta,q)$, so

$$s_q(\theta,q) = v_q(\theta,q) - c'(q) = \frac{\theta}{q} - c$$

and

$$\frac{s_q(\theta, q)}{v_{\theta q}(\theta, q)} = \frac{\theta\nu'(q) - c}{\nu'(q)} = \theta - \frac{c}{\nu'(q)} = \theta - cq.$$

Then, since since $f'(\theta) = 0$, $v_{\theta\theta q}(\cdot, \cdot) = 0$, and marginal cost is constant, (42) can be rewritten as

$$-\frac{\left[2 - (\overline{\theta} - \underline{\theta})\overline{\gamma}'(\theta)\right] v_{\theta q}(\theta, \overline{q}(\theta))}{v_{\theta q}(\theta, \overline{q}(\theta))v_{qq}(\theta, \overline{q}(\theta)) - s_q(\theta, \overline{q}(\theta))v_{\theta qq}(\theta, \overline{q}(\theta))} \geq -\frac{1}{v_{qq}(\theta, \overline{q}(\theta)) - \overline{c}''(\overline{q}(\theta))}$$

or, equivalently,

$$\frac{\left[2 - (\overline{\theta} - \underline{\theta})\overline{\gamma}'(\theta)\right] \frac{1}{\overline{q}(\theta)}}{\frac{c}{\overline{q}^2(\theta)}} \geq \frac{1}{\frac{\theta}{\overline{q}^2(\theta)} + \overline{c}''(\overline{q}(\theta))} \Leftrightarrow \frac{\left[2 - (\overline{\theta} - \underline{\theta})\overline{\gamma}'(\theta)\right] \overline{q}(\theta)}{c} \geq \frac{\overline{q}^2(\theta)}{\theta + \overline{q}^2(\theta)\overline{c}''(\overline{q}(\theta))}.$$

which, substituting the expression for $\overline{q}(\theta)$ and $\overline{\gamma}'(\theta) = 1/(\overline{\theta} - \underline{\theta})$, can be rewritten as

$$\frac{\left(\frac{\theta}{\underline{\theta}}\right)^i}{c} \geq \frac{i \exp\left(\frac{I + i\underline{\theta} + \overline{u}}{\underline{\theta}}\right)\left(\frac{\theta}{\underline{\theta}}\right)^{2i}}{\theta}$$

or, using the fact that $\underline{\theta} > 0$ by (IA5), as

$$1 \geq \frac{ci}{\underline{\theta}^i} \exp\left(\frac{I + i\underline{\theta} + \overline{u}}{\underline{\theta}}\right) \theta^{i-1}. \tag{43}$$

Condition (43) is certainly satisfied for c small enough for any given positive i. □

Construction of the Auxiliary Tariff when Income is Nonlinear in Type: Suppose the consumer's utility function is given by $v(\theta, q) = \theta\nu(q)$. In this case, the two conditions in (2) specialize to

$$\begin{cases} \overline{c}'(q) = \overline{\theta}(q)\nu'(q) \\ \overline{c}(q) = I + i(\overline{\theta}(q)). \end{cases}$$

From the second expression, it follows $\overline{c}'(q) = i'(\overline{\theta}(q))\overline{\theta}'(q)$ and substituting this latter expression into the first expression above, it follows

$$\frac{i'(\overline{\theta}(q))\overline{\theta}'(q)}{\overline{\theta}(q)} = \nu'(q). \tag{44}$$

Suppose, without loss, that $i(\overline{\theta}(q)) = i_1\overline{\theta}(q) + i_2\overline{\theta}^2(q) + i_3\overline{\theta}^3(q) + \cdots + i_n\overline{\theta}^n(q)$ for some n large enough. Then, (44) can be rewritten as

$$\left[\frac{i_1}{\overline{\theta}(q)} + 2i_2 + 3i_3\overline{\theta}(q) + \cdots + ni_n\overline{\theta}^{n-2}(q)\right]\overline{\theta}'(q) = \nu'(q).$$

Integrating both sides of this expression from $\overline{q}(\underline{\theta})$ to q, we obtain

$$\left[i_1 \log[\overline{\theta}(x)] + 2i_2\overline{\theta}(x) + \frac{3}{2}i_3\overline{\theta}^2(x) + \cdots + \frac{ni_n}{n-1}\overline{\theta}^{n-1}(x)\right]_{\overline{q}(\underline{\theta})}^{q} = [\nu(x)]_{\overline{q}(\underline{\theta})}^{q}$$

which implies

$$\begin{aligned} &i_1 \log[\overline{\theta}(q)] + 2i_2\overline{\theta}(q) + \cdots + \frac{ni_n}{n-1}\overline{\theta}^{n-1}(q) \\ -&i_1 \log[\overline{\theta}(\overline{q}(\underline{\theta}))] - 2i_2\overline{\theta}(\overline{q}(\underline{\theta})) - \cdots - \frac{ni_n}{n-1}\overline{\theta}^{n-1}(\overline{q}(\underline{\theta})) = \nu(q) - \nu(\overline{q}(\underline{\theta})) \end{aligned}$$

or, equivalently, by using $\overline{q}(\underline{\theta}) = (\nu)^{-1}\left(\frac{\overline{u}+I(\underline{\theta})}{\underline{\theta}}\right)$ and the fact that $\overline{\theta}(\overline{q}(\underline{\theta})) = \underline{\theta}$,

$$\begin{aligned} &i_1 \log[\overline{\theta}(q)] + 2i_2\overline{\theta}(q) + \cdots + \frac{ni_n}{n-1}\overline{\theta}^{n-1}(q) \\ &= \nu(q) - \frac{\overline{u}+I(\underline{\theta})}{\underline{\theta}} + i_1 \log(\underline{\theta}) + 2i_2\underline{\theta} + \cdots + \frac{ni_n}{n-1}\underline{\theta}^{n-1}. \end{aligned} \tag{45}$$

Since

$$i_1 \log[\overline{\theta}(q)] \approx i_1 \log[\overline{\theta}(q_0)] + \frac{i_1\left[\overline{\theta}(q) - \overline{\theta}(q_0)\right]}{\overline{\theta}(q_0)} - i_1\frac{\left[\overline{\theta}(q) - \overline{\theta}(q_0)\right]^2}{2\overline{\theta}^2(q_0)} + \cdots$$

it follows that, for given $\nu(q)$, as long as I is large enough and the slope of $I(\theta)$ is not too large or, alternatively, as long as I is not too large and the slope of $I(\theta)$ is sufficiently large, then $\overline{\theta}(q)$ is uniquely defined by (45). $\square$

Curvature of the Price Schedule: Recall that the seller's first-order condition for the optimal choice of quantity is

$$\sigma_q(\gamma,\theta,q) = v_q(\theta,q(\theta)) - c'(q(\theta)) - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}v_{\theta q}(\theta,q(\theta)) = 0$$

which can be rewritten as

$$v_q(\theta,q(\theta)) - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}v_{\theta q}(\theta,q(\theta)) = c'(q(\theta)) \Leftrightarrow \frac{v_q(\theta,q(\theta)) - c'(q(\theta))}{v_{\theta q}(\theta,q(\theta))} - \frac{\gamma(\theta)-F(\theta)}{f(\theta)} = 0 \tag{46}$$

or, equivalently, using the fact that $s_q(\theta,q(\theta)) = v_q(\theta,q(\theta)) - c'(q(\theta))$, as

$$\frac{s_q(\theta,q(\theta))}{v_{\theta q}(\theta,q(\theta))} = \frac{\gamma(\theta)-F(\theta)}{f(\theta)}.$$

So, by the implicit function theorem,

$$q'(\theta) = \frac{dq}{d\theta} = \frac{\frac{\partial}{\partial \theta}\left(\frac{s_q(\theta,q(\theta))}{v_{\theta q}(\theta,q(\theta))}\right) - \frac{\partial}{\partial \theta}\left(\frac{\gamma(\theta)-F(\theta)}{f(\theta)}\right)}{-\frac{\partial}{\partial q}\left(\frac{s_q(\theta,q(\theta))}{v_{\theta q}(\theta,q(\theta))}\right)} > 0$$

and

$$\gamma'(\theta) = \frac{d\gamma}{d\theta} = \frac{\frac{\partial}{\partial \theta}\left(\frac{s_q(\theta,q(\theta))}{v_{\theta q}(\theta,q(\theta))}\right) + \frac{\gamma(\theta)f'(\theta)}{f^2(\theta)} + \frac{\partial}{\partial \theta}\left(\frac{F(\theta)}{f(\theta)}\right)}{\frac{1}{f(\theta)}} > 0. \tag{47}$$

Suppose that $v(\theta, q) = \theta\nu(q)$ and that the seller's marginal cost is constant. Then, (46) can be rewritten as

$$\left[\theta + \frac{F(\theta)}{f(\theta)} - \frac{\gamma(\theta)}{f(\theta)}\right]\nu'(q(\theta)) = c \tag{48}$$

which also implies

$$\frac{\gamma(\theta)}{f(\theta)} = \theta + \frac{F(\theta)}{f(\theta)} - \frac{c}{\nu'(q(\theta))}. \tag{49}$$

Note that, with $0 \leq \gamma(\theta) \leq 1$, it follows

$$m(\theta) \equiv \theta + \frac{F(\theta) - \gamma(\theta)}{f(\theta)} \geq \theta - \frac{1 - F(\theta)}{f(\theta)} > 0$$

where the last inequality follows from the fact that

$$\theta - \frac{1 - F(\theta)}{f(\theta)} > 0 \Leftrightarrow 1 > \frac{1 - F(\theta)}{\theta f(\theta)}.$$

To see why $[1 - F(\theta)]/\theta f(\theta) < 1$, note that, by contradiction, if $[1 - F(\theta)]/\theta f(\theta) \geq 1$, then

$$\frac{T'(q) - c}{T'(q)} = \frac{1 - F(\theta)}{\theta f(\theta)} \geq 1$$

which would imply that at an optimal solution in the standard model, in which $\gamma(\theta) = 1$ at all θ's, $c \leq 0$. Hence, if $c > 0$, then $\theta - [1 - F(\theta)]/f(\theta) > 0$, so $m(\theta) > 0$.

Also, from (46) we obtain

$$q'(\theta) = \frac{dq}{d\theta} = \frac{m'(\theta)}{m(\theta)}\left[-\frac{\nu'(q(\theta))}{\nu''(q(\theta))}\right] \tag{50}$$

so for $q(\theta)$ to be increasing in θ (or $l(\gamma, \theta)$ to be increasing in θ for each γ), it must be $m'(\theta) > 0$, where

$$m'(\theta) = \frac{\partial}{\partial \theta}\left(\theta + \frac{F(\theta) - \gamma(\theta)}{f(\theta)}\right).$$

Note that from $T'(q) = \theta(q)\nu'(q)$ it follows that $T'(q) > 0$ an, by using (50),

$$T''(q) = \theta'(q)\nu'(q) + \theta(q)\nu''(q) = \frac{\nu'(q)}{q'(\theta)} + \theta(q)\nu''(q) = \frac{m(\theta)}{m'(\theta)}\left(-\nu''(q)\right) + \theta\nu''(q). \tag{51}$$

So,

$$T''(q) \leq 0 \Leftrightarrow -\nu''(q)\left[\frac{m(\theta)}{m'(\theta)} - \theta\right] \leq 0$$

which is trivially satisfied if $\nu''(q) = 0$. If, instead, $\nu''(q) < 0$, then

$$T''(q) \leq 0 \Leftrightarrow \frac{m(\theta)}{m'(\theta)} = \frac{\theta + \frac{F(\theta)}{f(\theta)} - \frac{\gamma(\theta)}{f(\theta)}}{\frac{\partial}{\partial\theta}\left(\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right)} \leq \theta$$

so a sufficient condition for $T''(q) \leq 0$ is

$$1+\frac{F(\theta)}{\theta f(\theta)} \leq \frac{\partial}{\partial\theta}\left(\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right) = 1+\frac{\partial}{\partial\theta}\left(\frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right) \leq 1+\frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right)+\min\left\{0, \frac{f'(\theta)}{f^2(\theta)}\right\}$$

since $0 \leq \gamma(\theta) \leq 1$. Similarly, since $\gamma(\theta) \leq 1$, $m(\theta) > 0$ and $m'(\theta) > 0$,

$$T''(q) = -\nu''(q)\left[\frac{\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}}{\frac{\partial}{\partial\theta}\left(\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right)} - \theta\right] \geq -\nu''(q)\left[\frac{\theta - \frac{1-F(\theta)}{f(\theta)}}{\frac{\partial}{\partial\theta}\left(\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right)} - \theta\right]$$

so a sufficient condition for $T''(q) \geq 0$ is

$$-\nu''(q)\left[\frac{\theta - \frac{1-F(\theta)}{f(\theta)}}{\frac{\partial}{\partial\theta}\left(\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right)} - \theta\right] \geq 0$$

and, if $\nu''(q) < 0$, a sufficient condition for $T''(q) \geq 0$ is

$$\theta - \frac{1-F(\theta)}{f(\theta)} \geq \theta\left[\frac{\partial}{\partial\theta}\left(\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right)\right] \Leftrightarrow -\frac{1-F(\theta)}{\theta f(\theta)} \geq \frac{\partial}{\partial\theta}\left(\frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right). \tag{52}$$

Lastly, note that from (48) it follows

$$T''(q) = -\nu''(q)\left[\frac{\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}}{\frac{\partial}{\partial\theta}\left(\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right)} - \theta\right] = \left[\frac{\frac{-\nu''(q)c}{\nu'(q)}}{\frac{\partial}{\partial\theta}\left(\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right)} - \theta\left(-\nu''(q)\right)\right]$$

so $T''(q) \le 0$ if, and only if,

$$\frac{c}{\theta}\left(-\frac{\nu''(q)}{\nu'(q)}\right) \le (-\nu''(q))\frac{\partial}{\partial\theta}\left(\theta + \frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right) = (-\nu''(q))\left[1+\frac{\partial}{\partial\theta}\left(\frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right)\right]. \tag{53}$$

Now, observe that (47) implies

$$\frac{\gamma'(\theta)}{f(\theta)} = \frac{\partial}{\partial\theta}\left(\frac{s_q(\theta,q(\theta))}{v_{\theta q}(\theta,q(\theta))}\right) + \frac{\gamma(\theta)f'(\theta)}{f^2(\theta)} + \frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right) = 1 - \frac{\partial}{\partial\theta}\left(\frac{c}{\nu'(q)}\right) + \frac{\gamma(\theta)f'(\theta)}{f^2(\theta)} + \frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right)$$

so when $\nu''(q) < 0$, (53) can be rewritten as

$$\frac{c}{\nu'(q)} \le \theta\left[1+\frac{\partial}{\partial\theta}\left(\frac{F(\theta)-\gamma(\theta)}{f(\theta)}\right)\right] = \theta\left[1+\frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right) - \frac{\gamma'(\theta)}{f(\theta)} + \frac{\gamma(\theta)f'(\theta)}{f^2(\theta)}\right]$$

or, equivalently, using (47) and (49),

$$\begin{aligned}\frac{c}{\nu'(q)} &\le \theta\left\{1+\frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right) - \left[1-\frac{\partial}{\partial\theta}\left(\frac{c}{\nu'(q)}\right) + \frac{\gamma(\theta)f'(\theta)}{f^2(\theta)} + \frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right)\right] + \frac{\gamma(\theta)f'(\theta)}{f^2(\theta)}\right\} \\ &= \theta\left[\frac{\partial}{\partial\theta}\left(\frac{c}{\nu'(q)}\right) - \frac{\gamma(\theta)f'(\theta)}{f^2(\theta)} + \frac{\gamma(\theta)f'(\theta)}{f^2(\theta)}\right] = \theta\left[\frac{\partial}{\partial\theta}\left(\frac{c}{\nu'(q)}\right)\right].\end{aligned}$$

From this we can conclude that when $\nu''(q) < 0$, then $T''(q) \ge 0$ if, and only if,

$$\frac{c}{\theta\nu'(q)} \le \frac{\partial}{\partial\theta}\left(\frac{c}{\nu'(q)}\right).$$

Lemma 14. *Assume that the seller's marginal cost is constant and $v(\theta,q) = \theta\nu(q)$. The price schedule $T(q)$ exhibits quantity discounts if $\nu''(q) \le 0$ and*

$$\frac{F(\theta)}{\theta f(\theta)} \le \frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right) + \min\left\{0, \frac{f'(\theta)}{f^2(\theta)}\right\}. \tag{54}$$

In general, when $\nu''(q) < 0$ the price schedule $T(q)$ exhibits quantity discounts if, and only if,

$$\frac{c}{\theta\nu'(q)} \le \frac{\partial}{\partial\theta}\left(\frac{c}{\nu'(q)}\right). \tag{55}$$

To understand the restriction under (54), assume that θ is uniformly distributed on $[\underline{\theta},\overline{\theta}]$, so $f(\theta) = 1/(\overline{\theta}-\underline{\theta})$ and $F(\theta) = (\theta-\underline{\theta})/(\overline{\theta}-\underline{\theta})$. Then, (54) is

$$\frac{(\theta-\underline{\theta})(\overline{\theta}-\underline{\theta})}{\theta(\overline{\theta}-\underline{\theta})} \le \frac{\partial}{\partial\theta}\left(\frac{(\theta-\underline{\theta})(\overline{\theta}-\underline{\theta})}{(\overline{\theta}-\underline{\theta})}\right) \Leftrightarrow 0 \le \frac{\underline{\theta}}{\theta}$$

which is certainly satisfied when $\underline{\theta} > 0$, as we assume.

Note that the curvature of the price schedule formally reflects the extent to which the price schedule displays 'bulk discounting', that is, unit prices declining in quantity. To see this, observe that, by definition, the unit price $p(q)$ is given by $p(q) = T(q)/q$, so unit prices decline in quantity if, and only if,

$$p'(q) = \frac{T'(q)q - T(q)}{q^2} \leq 0 \Leftrightarrow T'(q)q \leq T(q).$$

Observe now that if a function $f(x)$ is concave, then at any point $(x_2, f(x_2))$, $f(x) \leq f'(x_2)(x - x_2) + f(x_2)$, where $y(x) = f'(x_2)(x - x_2) + f(x_2)$ is the equation of the tangent line to $f(x)$ at $x = x_2$. This differential characterization of concavity requires continuous differentiability of $f(\cdot)$ (see Theorem M.C.1 in Mas-Colell, Whinston, and Green (1995)). Now, the condition $f(x) \leq f'(x_2)(x - x_2) + f(x_2)$ at $x = 0$ can be rewritten as

$$f'(x_2)x_2 \leq f(x_2) - f(0) \leq f(x_2)$$

where the last inequality holds if $f(0) \geq 0$. So, if f is concave and $f(0) \geq 0$, then $f'(x_2)x_2 \leq f(x_2)$ and $f(x_2)/x_2$ is declining in x_2. Interpret $f(\cdot)$ as $T(\cdot)$ and note that, without loss, we can rule out positive transfers from the consumer to the seller (we argued that such transfers can only occur between the seller and consumer types that do not trade, but in this case such transfers are not optimal for the seller). Thus, these observations immediately imply that if $T(\cdot)$ is concave, then unit prices decline in quantity.

B Appendix B

Proof of Lemma 12: From (26), $g(q) = f(\theta)/q'(\theta) = f(\theta)\theta'(q)$, and $\theta(q) = T'(q)/\nu'(q)$, it follows

$$T'(q) = c + \frac{[\gamma(\theta(q)) - F(\theta(q))]T'(q)}{\theta(q)f(\theta(q))} = c + \frac{\theta'(q)[\gamma(\theta(q)) - G(q)]v'(q)}{g(q)}$$

which implies

$$\theta'(q)v'(q) = \frac{g(q)\,[T'(q) - c]}{\gamma(\theta(q)) - G(q)}. \tag{56}$$

Differentiating $T'(q) = \theta(q)\nu'(q)$ with respect to q, we obtain

$$T''(q) = \theta'(q)\nu'(q) + \theta(q)\nu''(q) \Leftrightarrow \theta'(q)\nu'(q) = T''(q) - \theta(q)\nu''(q).$$

Substituting the resulting expression for $\theta'(q)\nu'(q)$ in (56), it follows

$$\theta(q)\nu''(q) = T''(q) - \frac{g(q)\,[T'(q) - c]}{\gamma(\theta(q)) - G(q)}.$$

By dividing the left-hand side of the above equality by $\theta(q)\nu'(q)$ and the right-hand side by $T'(q)$ (recall that $\theta(q)\nu'(q) = T'(q)$), we obtain

$$\frac{\nu''(q)}{\nu'(q)} = \frac{T''(q)}{T'(q)} - \frac{g(q)}{\gamma(\theta(q)) - G(q)}\left[1 - \frac{c}{T'(q)}\right]. \tag{57}$$

By integrating this expression from $q > \underline{q}$ to q_1, it follows

$$\log\left[\frac{\nu'(q_1)}{\nu'(q)}\right] = \log\left[\frac{T'(q_1)}{T'(q)}\right] - \int_q^{q_1} \frac{g(x)}{\gamma(\theta(x)) - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx.$$

Taking the exponential of this last expression and, then, dividing the left-hand side of it by $\theta_1\nu'(q_1)$ and the right-hand side by $T'(q_1)$ (recall that $\theta_1\nu'(q_1) = T'(q_1)$), it follows

$$\frac{1}{\theta_1\nu'(q)} = \frac{1}{T'(q)} \exp\left\{-\int_q^{q_1} \frac{g(x)}{\gamma(\theta(x)) - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx\right\}$$

or, equivalently,

$$\nu'(q) = \frac{T'(q)}{\theta_1} \exp\left\{-\int_q^{q_1} \frac{g(x)}{G(x)}\left[1 - \frac{c}{T'(x)}\right] dx\right\}$$

since $\gamma(\theta(x)) = 0$ for any $q \leq q_1$. Now, integrating (57) from q_2 to $q < \overline{q}$, it follows

$$\log\left[\frac{\nu'(q)}{\nu'(q_2)}\right] = \log\left[\frac{T'(q)}{T'(q_2)}\right] - \int_{q_2}^{q} \frac{g(x)}{\gamma(\theta(x)) - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx.$$

Taking the exponential of both sides of this latter equality and, then, multiplying the left-hand side of it by $\theta_2\nu'(q_2)$ and the right-hand side by $T'(q_2)$, we obtain

$$\nu'(q) = \frac{T'(q)}{\theta_2} \exp\left\{-\int_{q_2}^{q} \frac{g(x)}{1 - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx\right\}$$

since $\gamma(\theta(x)) = 1$ for any $q \geq q_2$. Similarly, integrating (57) from q_1 to $q < \widetilde{q}$, we obtain

$$\log\left[\frac{\nu'(q)}{\nu'(q_1)}\right] = \log\left[\frac{T'(q)}{T'(q_1)}\right] - \int_{q_1}^{q} \frac{g(x)}{\gamma(\theta(x)) - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx.$$

Taking the exponential of this last expression, multiplying the left-hand side of the above by $\theta_1\nu'(q_1)$ and the right-hand side by $T'(q_1)$, we obtain

$$\nu'(q) = \frac{T'(q)}{\theta_1} \exp\left\{-\int_{q_1}^{q} \frac{g(x)}{\gamma(\theta(x)) - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx\right\}.$$

Finally, integrating (57) from $q > \widetilde{q}$ to q_2, it follows

$$\log\left[\frac{\nu'(q_2)}{\nu'(q)}\right] = \log\left[\frac{T'(q_2)}{T'(q)}\right] - \int_q^{q_2} \frac{g(x)}{\gamma(\theta(x)) - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx.$$

Then, taking the exponential of this last expression, dividing the left-hand side of it by $\theta_2\nu'(q_2)$ and the right-hand side by $T'(q_2)$, we obtain

$$\frac{1}{\theta_2\nu'(q)} = \frac{1}{T'(q)}\exp\left\{-\int_q^{q_2} \frac{g(x)}{\gamma(\theta(x)) - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx\right\}$$

or, equivalently,

$$\nu'(q) = \frac{T'(q)}{\theta_2}\exp\left\{\int_q^{q_2} \frac{g(x)}{\gamma(\theta(x)) - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx\right\}.$$

□

Proof of Lemma 13: As stated in the main text, we only need to prove the second part of the result, related to the identifiability of $\nu'(q)$ and $\theta(q)$ and, thus, $f(\theta)$. Repeat the argument in the proof of Lemma 12 until expression (57) is derived. Integrating this expression from $\underline{q}$ to $q < \widetilde{q}$, it follows

$$\log\left[\frac{\nu'(q)}{\nu'(\underline{q})}\right] = \log\left[\frac{T'(q)}{T'(\underline{q})}\right] - \int_{\underline{q}}^{q} \frac{g(x)}{\gamma - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx.$$

Taking the exponential of both sides of this equality, and then multiplying the resulting left-hand side by $\underline{\theta}\nu'(\underline{q})$ and the resulting right-hand side by $T'(\underline{q})$, we obtain

$$\nu'(q) = \frac{T'(q)}{\underline{\theta}}\exp\left\{-\int_{\underline{q}}^{q} \frac{g(x)}{\gamma - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx\right\}.$$

Similarly, integrating (57) from $\widetilde{q} + \varepsilon$ to q with $\widetilde{q} + \varepsilon = q(\widetilde{\theta} + \varepsilon_\theta)$, where $\varepsilon, \varepsilon_\theta > 0$, it follows

$$\log\left[\frac{\nu'(q)}{\nu'(\widetilde{q} + \varepsilon)}\right] = \log\left[\frac{T'(q)}{T'(\widetilde{q} + \varepsilon)}\right] + \int_{\widetilde{q}+\varepsilon}^{q} \frac{g(x)}{G(x) - \gamma}\left[1 - \frac{c}{T'(x)}\right] dx.$$

Again, taking the exponential of the above expression, and then dividing the resulting left-hand side by $(\widetilde{\theta} + \varepsilon_\theta)\nu'(\widetilde{q} + \varepsilon)$ and the resulting right-hand side by $T'(\widetilde{q} + \varepsilon)$, we obtain

$$\nu'(q) = \frac{T'(q)}{(\widetilde{\theta} + \varepsilon_\theta)}\exp\left\{\int_{\widetilde{q}+\varepsilon}^{q} \frac{g(x)}{G(x) - \gamma}\left[1 - \frac{c}{T'(x)}\right] dx\right\}.$$

□

Alternative Estimators of Base Marginal Utility: Consider first the weakly-convex case.

When $q \in (\underline{q}, q_1)$ we can show that

$$\nu'(q) = \frac{T'(q)G(q)^{1-\frac{c}{T'(q)}}}{\theta_1 G(q_1)^{1-\frac{c}{T'(q_1)}}} \exp\left\{c \int_q^{q_1} \frac{\log\left[G(x)\right] T''(x)}{\left[T'(x)\right]^2} dx\right\}.$$

When $q \in (q_2, \overline{q})$ we can show that

$$\nu'(q) = \frac{T'(q)\left[1-G(q)\right]^{1-\frac{c}{T'(q)}}}{\theta_2 \left[1-G(q_2)\right]^{1-\frac{c}{T'(q_2)}}} \exp\left\{-c \int_{q_2}^{q} \frac{\log\left[1-G(x)\right] T''(x)}{\left[T'(x)\right]^2} dx\right\}.$$

When $q \in [q_1, q_2]$ such that $\gamma(\theta(q)) > G(q)$ or, equivalently, $G(q) > G(q_1)/[1 - G(q_2) + G(q_1)]$,

$$\nu'(q) = \frac{T'(q)\left\{G(q_2)\left[1-G(q_2)+G(q_1)\right]-G(q_1)\right\}^{\frac{G(q_2)-G(q_1)}{1-G(q_2)+G(q_1)}\left[1-\frac{c}{T'(q_2)}\right]}}{\theta_2\left\{G(q)\left[1-G(q_2)+G(q_1)\right]-G(q_1)\right\}^{\frac{G(q_2)-G(q_1)}{1-G(q_2)+G(q_1)}\left[1-\frac{c}{T'(q)}\right]}}$$

$$\cdot \exp\left(-\frac{c\left[G(q_2)-G(q_1)\right]}{1-G(q_2)+G(q_1)} \int_q^{q_2} \frac{\log\left\{G(x)\left[1-G(q_2)+G(q_1)\right]-G(q_1)\right\} T''(x)}{\left[T'(x)\right]^2} dx\right).$$

For all $q \in [q_1, q_2]$ such that $\gamma(\theta(q)) < G(q)$ or, equivalently, $G(q) < G(q_1)/[1 - G(q_2) + G(q_1)]$,

$$\nu'(q) = \frac{T'(q)\left\{G(q)\left[G(q_2)-G(q_1)-1\right]+G(q_1)\right\}^{\frac{G(q_2)-G(q_1)}{G(q_2)-G(q_1)-1}\left[1-\frac{c}{T'(q)}\right]}}{\theta_1\left\{G(q_1)\left[G(q_2)-G(q_1)-1\right]+G(q_1)\right\}^{\frac{G(q_2)-G(q_1)}{G(q_2)-G(q_1)-1}\left[1-\frac{c}{T'(q_1)}\right]}}$$

$$\cdot \exp\left(-\frac{c\left[G(q_2)-G(q_1)\right]}{G(q_2)-G(q_1)-1} \int_{q_1}^{q} \frac{\log\left\{G(x)\left[G(q_2)-G(q_1)-1\right]+G(q_1)\right\} T''(x)}{\left[T'(x)\right]^2} dx\right).$$

Recall that, for $q \in [q_1, q_2]$, $\gamma(q) = [G(q) - G(q_1)]/[G(q_2) - G(q_1)]$, so

$$\gamma(q) = \frac{G(q)-G(q_1)}{G(q_2)-G(q_1)} \gtreqless G(q) \Leftrightarrow G(q) \gtreqless \frac{G(q_1)}{1-G(q_2)+G(q_1)}.$$

First, note that for all q such that $\gamma(q) > G(q)$ or $G(q) > G(q_1)/[1 - G(q_2) + G(q_1)]$, ε_θ is such that $\widetilde{q} + \varepsilon = q(\widetilde{\theta} + \varepsilon_\theta)$, ε, $\varepsilon_\theta > 0$,

$$\begin{aligned}
\nu'(q) &= \frac{T'(q)}{(\widetilde{\theta}+\varepsilon_\theta)} \exp\left\{-\int_{\widetilde{q}+\varepsilon}^{q} \frac{g(x)}{\frac{G(x)-G(q_1)}{G(q_2)-G(q_1)} - G(x)} \left[1 - \frac{c}{T'(x)}\right] dx\right\} \\
&= \frac{T'(q)}{(\widetilde{\theta}+\varepsilon_\theta)} \exp\left\{-\frac{\left[G(q_2)-G(q_1)\right]}{1-G(q_2)+G(q_1)} \int_{\widetilde{q}+\varepsilon}^{q} \frac{[1-G(q_2)+G(q_1)]g(x)}{[1-G(q_2)+G(q_1)]G(x)-G(q_1)} dx\right\} \\
&\quad \cdot \exp\left\{\frac{G(q_2)-G(q_1)}{1-G(q_2)+G(q_1)} \int_{\widetilde{q}+\varepsilon}^{q} \frac{[1-G(q_2)+G(q_1)]g(x)}{\left[1-G(q_2)+G(q_1)\right]G(x)-G(q_1)} \cdot \frac{c}{T'(x)} dx\right\}
\end{aligned}$$

or, equivalently

$$
\begin{aligned}
\nu'(q) \;\; = \;\; & \frac{T'(q)}{(\widetilde{\theta}+\varepsilon_\theta)} \exp\left\{-\frac{[G(q_2)-G(q_1)]}{[1-G(q_2)+G(q_1)]}\left[\log\left\{[1-G(q_2)+G(q_1)]G(x)-G(q_1)\right\}\right]_{\widetilde{q}+\varepsilon}^{q}\right\} \\
& \cdot \exp\left\{\frac{c[G(q_2)-G(q_1)]}{1-G(q_2)+G(q_1)}\left[\frac{\log\left\{[1-G(q_2)+G(q_1)]G(x)-G(q_1)\right\}}{T'(x)}\right]_{\widetilde{q}+\varepsilon}^{q}\right\} \\
& \cdot \exp\left\{\frac{c[G(q_2)-G(q_1)]}{1-G(q_2)+G(q_1)}\int_{\widetilde{q}+\varepsilon}^{q}\frac{\log\left\{[1-G(q_2)+G(q_1)]G(x)-G(q_1)\right\}T''(x)}{[T'(x)]^2}\right\}
\end{aligned}
$$

or, equivalently

$$
\begin{aligned}
\nu'(q) \;\; = \;\; & \frac{T'(q)}{(\widetilde{\theta}+\varepsilon_\theta)} \cdot \frac{\left\{[1-G(q_2)+G(q_1)]G(q)-G(q_1)\right\}^{-\frac{G(q_2)-G(q_1)}{[1-G(q_2)+G(q_1)]}}}{\left\{[1-G(q_2)+G(q_1)]G(\widetilde{q}+\varepsilon)-G(q_1)\right\}^{-\frac{G(q_2)-G(q_1)}{[1-G(q_2)+G(q_1)]}}} \\
& \cdot \exp\left\{\frac{c[G(q_2)-G(q_1)]}{[1-G(q_2)+G(q_1)]}\left[\log\left(\left\{[1-G(q_2)+G(q_1)]G(x)-G(q_1)\right\}^{\frac{1}{T'(x)}}\right)\right]_{\widetilde{q}+\varepsilon}^{q}\right\} \\
& \cdot \exp\left\{\frac{c[G(q_2)-G(q_1)]}{[1-G(q_2)+G(q_1)]}\int_{\widetilde{q}+\varepsilon}^{q}\frac{\log\left\{[1-G(q_2)+G(q_1)]G(x)-G(q_1)\right\}T''(x)}{[T'(x)]^2}\right\}
\end{aligned}
$$

or, equivalently

$$
\begin{aligned}
\nu'(q) \;\; = \;\; & \frac{T'(q)}{(\widetilde{\theta}+\varepsilon_\theta)} \cdot \frac{\left\{[1-G(q_2)+G(q_1)]G(q)-G(q_1)\right\}^{-\frac{G(q_2)-G(q_1)}{1-G(q_2)+G(q_1)}}}{\left\{[1-G(q_2)+G(q_1)]G(\widetilde{q}+\varepsilon)-G(q_1)\right\}^{-\frac{G(q_2)-G(q_1)}{1-G(q_2)+G(q_1)}}} \\
& \cdot \frac{\left\{[1-G(q_2)+G(q_1)]G(q)-G(q_1)\right\}^{\frac{c[G(q_2)-G(q_1)]}{T'(q)[1-G(q_2)+G(q_1)]}}}{\left\{[1-G(q_2)+G(q_1)]G(\widetilde{q}+\varepsilon)-G(q_1)\right\}^{\frac{c[G(q_2)-G(q_1)]}{T'(\widetilde{q}+\varepsilon)[1-G(q_2)+G(q_1)]}}} \\
& \cdot \exp\left\{\frac{c[G(q_2)-G(q_1)]}{[1-G(q_2)+G(q_1)]}\int_{\widetilde{q}+\varepsilon}^{q}\frac{\log\left\{[1-G(q_2)+G(q_1)]G(x)-G(q_1)\right\}T''(x)}{[T'(x)]^2}\right\}
\end{aligned}
$$

or, equivalently

$$
\begin{aligned}
\nu'(q) \;\; = \;\; & \frac{T'(q)\left\{[1-G(q_2)+G(q_1)]G(q)-G(q_1)\right\}^{\frac{G(q_2)-G(q_1)}{1-G(q_2)+G(q_1)}\left[\frac{c}{T'(q)}-1\right]}}{(\widetilde{\theta}+\varepsilon_\theta)\left\{[1-G(q_2)+G(q_1)]G(\widetilde{q}+\varepsilon)-G(q_1)\right\}^{\frac{G(q_2)-G(q_1)}{1-G(q_2)+G(q_1)}\left[\frac{c}{T'(\widetilde{q}+\varepsilon)}-1\right]}} \\
& \cdot \exp\left\{\frac{c[G(q_2)-G(q_1)]}{1-G(q_2)+G(q_1)}\int_{\widetilde{q}+\varepsilon}^{q}\frac{\log\left\{[1-G(q_2)+G(q_1)]G(x)-G(q_1)\right\}T''(x)}{[T'(x)]^2}\right\}.
\end{aligned}
$$

Consider now the highly-convex case. For all q and $\gamma > G(q)$, with $q \in [\underline{q}, \widetilde{q})$ it follows

$$
\nu'(q) = \frac{T'(q)\left[\gamma - G(q)\right]^{1-\frac{c}{T'(q)}}}{\underline{\theta}\gamma^{1-\frac{c}{T'(\underline{q})}}} \exp\left\{-c\int_{\underline{q}}^{q}\frac{\log\left[\gamma - G(x)\right]T''(x)}{\left[T'(x)\right]^2}dx\right\}.
$$

For quantities in $[\widetilde{q}+\varepsilon, \overline{q})$, we have $G(x) > \gamma$ and, with ε_θ is such that $\widetilde{q}+\varepsilon = q(\widetilde{\theta}+\varepsilon_\theta)$, ε, $\varepsilon_\theta > 0$, that

$$\nu'(q) = \frac{T'(q)}{(\widetilde{\theta}+\varepsilon_\theta)} \exp\left\{-\int_{\widetilde{q}+\varepsilon}^{q} \frac{g(x)}{\gamma - G(x)} \left[1 - \frac{c}{T'(x)}\right] dx\right\}$$

or, equivalently,

$$\begin{aligned} \nu'(q) \quad = \quad & \frac{T'(q)}{(\widetilde{\theta}+\varepsilon_\theta)} \exp\left\{[\log[G(x)-\gamma]]_{\widetilde{q}+\varepsilon}^{q}\right\} \exp\left\{-c\left[\frac{\log[G(x)-\gamma]}{T'(x)}\right]_{\widetilde{q}+\varepsilon}^{q}\right. \\ & \left. \cdot -c \int_{\widetilde{q}+\varepsilon}^{q} \frac{\log[G(x)-\gamma]T''(x)}{[T'(x)]^2} dx\right\} \end{aligned}$$

or, equivalently,

$$\nu'(q) = \frac{T'(q)}{(\widetilde{\theta}+\varepsilon_\theta)} \frac{[G(q)-\gamma]^{1-\frac{c}{T'(q)}}}{[G(\widetilde{q}+\varepsilon)-\gamma]^{1-\frac{c}{T'(\widetilde{q}+\varepsilon)}}} \exp\left\{-c \int_{\widetilde{q}+\varepsilon}^{q} \frac{\log[G(x)-\gamma]T''(x)}{[T'(x)]^2} dx\right\}.$$

In each case, the corresponding expression for $\theta(q)$ is obtained from $\theta(q) = T'(q)/\nu'(q)$. $\square$

C Appendix C: Nonparametric Estimation

We start by constructing the price schedule in each village (municipality) with at least 200 observations on quantity consumed of rice.

Price Schedule. In our data we have direct information on unit prices and quantities purchased by each consumer (household) in each market (village). Then, the price schedule can be simply obtained by fitting, for instance by least squares, observed household expenditures, $T(q)$, on observed quantities, where $T(q) = p(q)q$, $p(q)$ denotes the unit value or observed price for quantity q, and q the observed quantity. (We restrict attention to quantities between 0 and 3 kilos consumed per week and sold at a (median) price per unit not higher than 12 pesos, approximately one dollar (that is, $0.912 = (0.0760 \cdot 12)$ dollars).) We consider six different specifications: (1) $T(q) = t_0 + t_1 q + t_2 q^2$; (2) $\log(T(q)) = t_0 + t_1 \log q + t_2 (\log q)^2$; (3) $\log(T(q)) = t_0 + t_1 \log q + t_2 (\log q)^2 + t_3 q$; (4) $\log(T(q)) = t_0 + t_1 \log(q)$; (5) $T(q) = t_0 + t_1 \log(q)$; and (6) $T(q) = \log(t_0 + t_1 q)$. The reason the quantity q in all these specifications can be treated as exogenous is that having information on the quantity purchased by each consumer is equivalent to having information on the price schedule of the seller. Also, according to the model, $T(q)$ is a deterministic function. If the fit is good, then the measurement error fitting may cause can be considered minimal and, thus, ignored.

For simplicity, we omit here the notation for the fact that the price schedule is estimated. Note that for the first specification, *Specification 1*,

$$T'(q) = t_1 + 2t_2 q$$

and $T''(q) = 2t_2$. For the second specification, *Specification 2*,

$$T(q) = \exp\{t_0 + t_1 \log q + t_2(\log q)^2\}$$

so

$$T'(q) = \exp\{t_0 + t_1 \log q + t_2(\log q)^2\} \left(\frac{t_1}{q} + \frac{2t_2 \log q}{q}\right) = T(q) \left(\frac{t_1 + 2t_2 \log q}{q}\right)$$

and

$$T''(q) = T(q) \cdot \frac{(t_1 + 2t_2 \log q)(t_1 + 2t_2 \log q - 1) + 2t_2}{q^2}.$$

For the third specification, *Specification 3*,

$$T(q) = \exp\{t_0 + t_1 \log q + t_2(\log q)^2 + t_3 q\}$$

so

$$T'(q) = \exp\{t_0 + t_1 \log q + t_2(\log q)^2 + t_3 q\} \left(\frac{t_1}{q} + \frac{2t_2 \log q}{q} + t_3\right) = T(q) \left(\frac{t_1 + 2t_2 \log q}{q} + t_3\right)$$

and

$$T''(q) = T(q) \left[\frac{(t_1 + 2t_2 \log q + t_3 q)(t_1 + 2t_2 \log q + t_3 q - 1) + 2t_2 + t_3 q}{q^2}\right].$$

For the fourth specification, *Specification 4*, we consider

$$\log(T(q)) = t_0 + t_1 \log(q)$$

or $T(q) = \exp\{t_0 + t_1 \log(q)\}$ with

$$T'(q) = \exp\{t_0 + t_1 \log(q)\} \frac{t_1}{q} = T(q) \cdot \frac{t_1}{q}$$

and

$$T''(q) = T'(q) \cdot \frac{t_1}{q} + T(q) \cdot \frac{-t_1}{q^2} = T(q) \cdot \frac{t_1(t_1 - 1)}{q^2}.$$

For the fifth specification, *Specification 5*, we consider

$$T(q) = t_0 + t_1 \log(q)$$

with $T'(q) = t_1/q$ and $T''(q) = -t_1/q^2$. Finally, for the sixth *Specification 6*, we consider

$$T(q) = \log(t_0 + t_1 q)$$

with $T'(q) = t_1/(t_0 + t_1 q)$ and $T''(q) = -t_1^2/(t_0 + t_1 q)^2$.

C.1 Standard Model

Recall that in this case $\gamma(\theta) = 1$ for all θ, where $\gamma(\theta)$ is the multiplier associated with the participation constraint/budget constraint of type θ. The standard model's primitives are $\{\nu(q), F(\theta), c\}$, where $\nu(q)$ denotes the 'base' utility function of the consumer (recall that $v(\theta, q) = \theta\nu(q)$), $F(\theta)$ denotes the cumulative distribution function of types, and c the seller's marginal cost, which we assume constant. From the necessary (and sufficient) conditions for optimality, that is,

$$\frac{T'(q) - c}{T'(q)} = \frac{1 - F(\theta)}{\theta f(\theta)} \tag{58}$$

it follows that $T'(\overline{q}) = c$, since $F(\overline{\theta}) = 1$. So c is identified by the price of the *largest* quantity observed purchased in each village. Next, using integration by parts and by straightforward algebraic manipulations of (58), it also follows

$$\nu'(q) = \frac{T'(q)}{\underline{\theta}} \exp\left\{-\int_{\underline{q}}^{q} \frac{g(x)}{1 - G(x)}\left[1 - \frac{c}{T'(x)}\right] dx\right\}$$

and, through further manipulation,

$$\nu'(q) = \frac{T'(q)\left[1 - G(q)\right]^{1 - \frac{c}{T'(q)}}}{\underline{\theta}} \exp\left\{-c\int_{\underline{q}}^{q} \frac{\log\left[1 - G(x)\right] T''(x)}{\left[T'(x)\right]^2} dx\right\}. \tag{59}$$

Note that $\nu'(\cdot)$ is a known function of observables, known, or estimable functions ($T'(q)$, $T''(q)$, $G(q)$, c, $\underline{\theta}$, and $\underline{q}$). Then, types can be generated from the condition for local incentive compatibility, $\theta = T'(q)/\nu'(q)$, with $\nu'(q)$ computed as in (59). The density $f(\theta)$ can be nonparametrically estimated from this pseudo-sample of types. Correspondingly, to estimate $\{\nu(q), f(\theta), c\}$ we proceed in a series of steps.

C.1.1 Estimation of Primitives

To recover the model's primitives, we proceed according to the following seven steps.

Step 1. Set $\underline{\theta} = 1$, since it is not identified.

Step 2. Compute $G(q_i)$, $T'(q_i)$, and $T''(q_i)$, $i = 1, \ldots, N$ from data on quantity purchases and prices (unit values) in each village. In particular,

$$\widehat{G}(q) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}(q_i \leq q)$$

where for $q \in [\underline{q}, \overline{q}]$ and $\mathbf{1}(\cdot)$ is an indicator function.

Step 3. Recover c as

$$\widehat{c} = T'(\overline{q})$$

where $\overline{q} = \max_i\{q_i\} = \max\{q_1, \ldots, q_N\}$ from the fact that $\overline{q} = q(\overline{\theta})$ and $[T'(q) - c]/T'(q) =$

$[1 - F(\theta)]/\theta f(\theta)$.

Step 4. Let $\underline{q} = \min_i\{q_i\} = \min\{q_1, \ldots, q_N\}$. Note that because the empirical distribution function of quantity purchases is a step function with steps at $q_1 < \cdots < q_N$, the integral from (59) can be rewritten as the finite sum of integrals from q_i to q_{i+1}, $i = 1, \ldots, N-1$. (See PV (2010) on this.) On each of these intervals $\log[1 - G(\cdot)]$ is constant whereas the primitive of $T''(\cdot)/[T'(\cdot)]^2$ is $-1/T'(\cdot)$. Let $q \in [\underline{q}, \overline{q})$. We compute

$$\widehat{\nu'}(q) = T'(q)\,[1 - G(q_i)]^{1 - \frac{\widehat{c}}{T'(q)}} \exp\left\{ -\widehat{c} \sum_{j=1}^{i-1} \log[1 - G(q_j)] \left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})} \right] \right.$$
$$\left. -\widehat{c} \log[1 - G(q_i)] \left[\frac{1}{T'(q_i)} - \frac{1}{T'(q)} \right] \right\} \tag{60}$$

with $q \in [q_i, q_{i+1})$, $i = 1, \ldots, N-1$, $q_1 = \underline{q}$, and $q_N = \overline{q}$. When $i = 1$, this expression reduces to

$$\widehat{\nu'}(q_1) = T'(q_1)\,[1 - G(q_1)]^{1 - \frac{\widehat{c}}{T'(q_1)}}.$$

When $q_N = \overline{q}$, compute marginal utility as

$$\lim_{q \uparrow q_N} \widehat{\nu'}(q) = T'(q_N)\,[1 - G(q_{N-1})]^{1 - \frac{\widehat{c}}{T'(q_N)}} \exp\left\{ -\widehat{c} \sum_{j=1}^{N-1} \log[1 - G(q_j)] \left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})} \right] \right\}.$$

Step 5. Compute θ_i, $i = 1, \ldots, N$, as

$$\widehat{\theta}_i = T'(q_i)/\widehat{\nu'}(q_i).$$

Note that we obtain $\widehat{\theta}_{N-1} = \widehat{\theta}_N$.

Step 6. Compute

$$\hat{f}_h(\theta) = \frac{1}{N} \sum_{i=1}^{N} K_h\left(\frac{\theta - \widehat{\theta}_i}{h} \right) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{\theta - \widehat{\theta}_i}{h} \right)$$

$i = 1, \ldots, N-1$, for a suitable choice of (symmetric) kernel function $K(\cdot)$ and bandwidth h ($K_h(x) = 1/hK(x/h)$ is a scaled kernel). (Observe that we do not need to compute this step in Fortran, given that in the calculations to follow we just rely on $F(\theta)$, which we know satisfies $F(\theta) = G(q)$.)

Step 7. Compute consumer surplus under nonlinear pricing as $CS_{np}(q) = \theta(q)\nu(q) - T(q)$, that is,

$$CS_{np} = \int_{\underline{q}}^{\overline{q}} CS_{np}(x)dG(x) = \int_{\underline{q}}^{\overline{q}} [\theta(x)\nu(x) - T(x)]dG(x).$$

Recall that at the optimum the participation constraint of the lowest type must bind. We also maintain that $\overline{u} = 0$. So, $\underline{\theta}\nu(\underline{q}) - T(\underline{q}) = 0$ and, with $\underline{\theta}$ normalized at one, it follows that

$\nu(\underline{q}) = T(\underline{q})$. Hence,

$$\nu(q) = \nu(\underline{q}) + \int_{\underline{q}}^{q} \nu'(x)dx = T(\underline{q}) + \int_{\underline{q}}^{q} \nu'(x)dx.$$

Using the fact that all relevant variables are discrete, we compute for $i = 1$,

$$\widehat{\nu}(q_1) = T(\underline{q}) + \widehat{\nu}(q_1) = T(\underline{q}) + q_1\widehat{\nu'}(q_1)$$

and for $i > 1$,

$$\widehat{\nu}(q_i) = T(\underline{q}) + q_1\widehat{\nu'}(q_1) + \sum_{j=1}^{i-1}(q_{j+1} - q_j)\widehat{\nu'}(q_{j+1}). \tag{61}$$

Specifically, note that

$$\begin{aligned} \widehat{\nu}(q_2) &= T(\underline{q}) + q_1\widehat{\nu'}(q_1) + \sum_{j=1}^{1}(q_{j+1} - q_j)\widehat{\nu'}(q_{j+1}) = T(\underline{q}) + q_1\widehat{\nu'}(q_1) + (q_2 - q_1)\widehat{\nu'}(q_2) \\ &= \widehat{\nu}(q_1) + (q_2 - q_1)\widehat{\nu'}(q_2) \end{aligned}$$

and

$$\begin{aligned} \widehat{\nu}(q_3) &= T(\underline{q}) + q_1\widehat{\nu'}(q_1) + \sum_{j=1}^{2}(q_{j+1} - q_j)\widehat{\nu'}(q_{j+1}) \\ &= T(\underline{q}) + q_1\widehat{\nu'}(q_1) + (q_2 - q_1)\widehat{\nu'}(q_2) + (q_3 - q_2)\widehat{\nu'}(q_3) = \widehat{\nu}(q_2) + (q_3 - q_2)\widehat{\nu'}(q_3). \end{aligned}$$

Therefore, for $i = 1$, we obtain

$$\widehat{\nu}(q_1) = T(\underline{q}) + q_1\widehat{\nu'}(q_1)$$

and for $i > 1$, we obtain

$$\widehat{\nu}(q_i) = \widehat{\nu}(q_{i-1}) + (q_i - q_{i-1})\widehat{\nu'}(q_i).$$

We compute consumer surplus as

$$\widehat{CS}_{np} = \sum_{i=1}^{N}[\widehat{\theta}(q_i)\widehat{\nu}(q_i) - T(q_i)]r_q(q_i)$$

where $r_q(q_1) = G(q_1)$ and, for $i = 1, \ldots, N-1$, $r_q(q_{i+1}) = G(q_{i+1}) - G(q_i)$. Similarly, we compute producer surplus as

$$\widehat{PS}_{np} = \sum_{i=1}^{N}\left[T(q_i) - cq_i\right]r_q(q_i)$$

given our maintained assumption that the seller's marginal cost is constant.

C.1.2 Counterfactual Experiment: Linear Pricing Outcomes

Here we compare consumer and producer surplus under standard nonlinear pricing and under (monopoly) linear pricing. To this purpose, we need to compute the seller's linear price, the corresponding individual quantity demanded by each type, and the aggregate quantity demanded by all types. We start by computing the quantity $q(p_m, \theta)$ demanded by a consumer of type θ as a function of the linear monopoly price p_m,

$$\theta\nu'(q) = p_m \Rightarrow q = q(p_m, \theta) = (\nu')^{-1}\left(\frac{p_m}{\theta}\right). \tag{62}$$

Let $Q(p_m) = \int_{\underline{\theta}}^{\overline{\theta}} q(p_m, x) f(x)dx$ be the corresponding aggregate quantity demanded at price p_m. Then, p_m solves the problem:

$$\max_{p_m}[(p_m - c)Q(p_m)]$$

with F.O.C. for the optimal choice of p_m given by

$$Q(p_m) + (p_m - c)Q'(p_m) = 0. \tag{63}$$

In particular,

$$\frac{p_m - c}{p_m} = -\frac{Q(p_m)}{p_m Q'(p_m)} = -\frac{1}{|\varepsilon_{p_m}^Q|}$$

where $\varepsilon_{p_m}^Q$ is the price elasticisty of aggregate demand. Using the definition of $Q(p_m)$ and (62), it follows

$$Q'(p_m) = \frac{\partial Q(p_m)}{\partial p_m} = \int_{\underline{\theta}}^{\overline{\theta}} \frac{\partial q(p_m, x)}{\partial p_m} f(x)dx = \int_{\underline{\theta}}^{\overline{\theta}} \frac{f(x)dx}{\nu''(q(p_m, x))x}$$

where the last equality in the above follows by applying the implicit function theorem to (62). Note that

$$-\frac{Q(p_m)}{p_m Q'(p_m)} = \frac{\int_{\underline{\theta}}^{\overline{\theta}} (\nu')^{-1}\left(\frac{p_m}{x}\right) f(x)dx}{\theta \int_{\underline{\theta}}^{\overline{\theta}} \frac{f(x)dx}{(-\nu''(q(p_m,x)))x}} \geq \frac{1 - F(\theta)}{\theta f(\theta)}$$

if, and only if,

$$\int_{\underline{\theta}}^{\overline{\theta}} (\nu')^{-1}\left(\frac{p_m}{x}\right) f(x)dx \geq \frac{1 - F(\theta)}{f(\theta)} \int_{\underline{\theta}}^{\overline{\theta}} \frac{f(x)dx}{(-\nu''(q(p_m, x)))x}$$

and at $\underline{\theta}$ this becomes

$$\int_{\underline{\theta}}^{\overline{\theta}} \left[f(\underline{\theta})(\nu')^{-1}\left(\frac{p_m}{x}\right) - \frac{1}{(-\nu''(q(p_m, x)))x}\right] f(x)dx \geq 0.$$

Recall our normalization $\underline{\theta} = 1$. Then the marginal price price charged to the lowest type is smaller than the monopoly price if $f(\underline{\theta})$ is large enough.

Hence, consumer surplus under linear pricing is given by

$$CS_{lp} = \int_{\underline{\theta}}^{\overline{\theta}} CS_{lp}(x)f(x)dx = \int_{\underline{\theta}}^{\overline{\theta}} \left[x\nu(q(p_m, x)) - p_m q(p_m, x)\right] f(x)dx$$

and producer surplus by

$$PS_{lp} = (p_m - c)Q(p_m) = (p_m - c)\int_{\underline{\theta}}^{\overline{\theta}} q(p_m, x)f(x)dx.$$

In practice, we proceed as follows:

(a) Determine a grid of values for p_m and denote it by $\mathbf{p}_m = (p_{m1}, \ldots, p_{mP})$;

(b) For each θ_i, determine the quantity demanded by each type for each possible price in $\mathbf{p}_m$. Denote the array of such quantities by $q(\mathbf{p}_m, \widehat{\theta}_i)$, where

$$q(\mathbf{p}_m, \widehat{\theta}_i) = q((p_{m1}, \ldots, p_{mP}), \widehat{\theta}_i) = (q(p_{m1}, \widehat{\theta}_i), \ldots, q(p_{mP}, \widehat{\theta}_i))$$

and for each type $\widehat{\theta}_i$, each quantity $q(p_{mp}, \widehat{\theta}_i)$ satisfies $\widehat{\theta}_i\nu'(q) = p_{m1p}$, $p = 1, \ldots, P$. In practice, to determine the quantity chosen by type $\widehat{\theta}_i$ for each possible price p_{mp}, $p = 1, \ldots, P$; we solve the system

$$\begin{cases} \widehat{\theta}_i\nu'(q_1) = p_{m1} \\ \ldots \\ \widehat{\theta}_i\nu'(q_N) = p_{m1} \\ \ldots \\ \widehat{\theta}_i\nu'(q_1) = p_{mP} \\ \ldots \\ \widehat{\theta}_i\nu'(q_N) = p_{mP} \end{cases}$$

where $\mathbf{q} = (q_1, \ldots, q_N)$ is an array of candidate quantities (we chose a grid of 1,000 equidistant points). For each village, we determine this grid as ranging from the minimum and maximum observed quantity;

(c) Compute p_m as follows. Let

$$Q(p_m) = (Q(p_{m1}), \ldots, Q(p_{mP}))$$

where, for $p = 1, \ldots, P$,

$$Q(p_{mp}) = \sum_{i=1}^{N} q(p_{mp}, \widehat{\theta}_i)r_\theta(\widehat{\theta}_i) = \sum_{i=1}^{N} q(p_{mp}, \widehat{\theta}_i)r_q(q_i),$$

and $r_\theta(\widehat{\theta}_1) = r_q(q_1) = G(q_1)$ whereas $r_\theta(\widehat{\theta}_{i+1}) = r_q(q_{i+1}) = G(q_{i+1}) - G(q_i)$, $i = 1, \ldots, N-1$. (This is due to the fact that, as discussed, the model implies $G(q_i) = F(\theta_i)$, so there are as many 'exactly' estimated types as observed quantities.) Then, for $p = 1, \ldots, P-1$ we compute

$$Q'(p_{mp}) \approx \frac{Q(p_{mp}+h) - Q(p_{mp})}{h} \approx \frac{Q(p_{mp+1}) - Q(p_{mp})}{p_{mp+1} - p_{mp}}$$

and for $p = P$ we compute

$$Q'(p_{mP}) \approx \frac{Q(p_{mP}) - Q(p_{mP}-h)}{h} \approx \frac{Q(p_{mP}) - Q(p_{mP-1})}{p_{mP} - p_{mP-1}}.$$

Lastly, we determine the linear monopoly price as the price p_m^* that solves the seller's first-order condition (63). In practice, we compute p_m^* as

$$p_m^* = \min_{p \in \{p_{m1}, \ldots, p_{mP}\}} \left\{ |Q(p) + (p-c)Q'(p)| \right\};$$

(d) From the above computation of the base utility function $\widehat{\nu}(\cdot)$, interpolating it over quantities not in $\{q_1, \ldots, q_N\}$, we can immediately compute $\widehat{\theta}_i \widehat{\nu}(\cdot)$ at each $q(p_m^*, \widehat{\theta}_i)$. We then calculate $\widehat{CS}_{lp}$ as

$$\widehat{CS}_{lp} = \sum_{i=1}^{N} [\widehat{\theta}_i \widehat{\nu}(q(p_m^*, \widehat{\theta}_i)) - p_m^* q(p_m^*, \widehat{\theta}_i)] r_\theta(\widehat{\theta}_i) = \sum_{i=1}^{N} \widehat{\theta}_i \widehat{\nu}(q(p_m^*, \widehat{\theta}_i)) r_q(q(p_m^*, \widehat{\theta}_i)) - p_m^* Q(p_m^*).$$

Lastly, since $r_\theta(\widehat{\theta}_i) = r_q(q(p_m^*, \widehat{\theta}_i))$ by $G(q_i) = F(\theta_i)$ and

$$Q(p_m^*) = \sum_{i=1}^{N} q(p_m^*, \widehat{\theta}_i) r_q(q(p_m^*, \widehat{\theta}_i)),$$

We compute producer surplus as

$$\widehat{PS}_{lp} = \sum_{i=1}^{N} (p_m^* - c)\, q(p_m^*, \widehat{\theta}_i) r_q(q(p_m^*, \widehat{\theta}_i)) = (p_m^* - c) Q(p_m^*).$$

C.2 Model with Type Dependent Reservation Utility

Consider now our model, in which, due to outside consumption possibilities or budget constraints that depend on the consumer's type, the consumer's reservation utility depends on θ, so the multiplier associated with the participation constraint, $\gamma(\theta)$, is not necessarily equal to one. Following our characterization of the optimal price schedule, we distinguish between the weakly-convex and the highly-convex case. We start with the highly-convex case, given that it represents a generalization of the standard model.

Result 1. *If $T'(\underline{q}) = T'(\overline{q})$, then observed prices and quantities correspond to the weakly-convex case. Otherwise, the highly-convex case applies.*

C.2.1 Highly-Convex Case

Suppose the highly-convex case applies. We proceed according to the following steps.

Step 1. Compute $\widehat{G}(q_i)$, $T'(q_i)$, and $T''(q_i)$, $i = 1, \ldots, N$ and $q_i \in \{q_1, \ldots, q_N\}$, from data on quantity purchases and prices (unit values) in each village. In particular, compute

$$\widehat{G}(q) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(q_i \leq q)$$

where $q \in [\underline{q}, \overline{q}]$ and $\mathbf{1}(\cdot)$ is an indicator function.

Step 2. Recall that $q_1 = \underline{q}$, and $q_N = \overline{q}$. Note that when $\gamma = 1$, we have $T'(q_i) \geq T'(q_N)$ for all $q_i \in \{q_1, \ldots, q_N\}$ and $c = T'(q_N)$. When $\gamma = 0$, we have $T'(q_i) \leq T'(q_1)$ for all $q_i \in \{q_1, \ldots, q_N\}$ and $c = T'(q_1)$. In each village we recover c as the estimated constant $\widehat{\beta}_0$ of the nonparametric regression discussed in the body of the paper.

Step 3. From the fact that $G(q) = \gamma - \theta q'(\theta)[T'(q) - c]g(q)/T'(q)$, compute γ as detailed in the main body of the paper.

Step 4. Recall that because the empirical distribution function of quantity purchases is a step function with steps at $q_1 < \cdots < q_N$, the integral in $\nu'(q)$ can be rewritten as the finite sum of integrals. (See Perrigne and Vuong (2010) on this.) On each of these intervals $\widehat{G}(\cdot)$ is constant whereas the integral of $T''(\cdot)/[T'(\cdot)]^2$ is $-1/T'(\cdot)$. We distinguish three cases.

Case 1 $(0 < \gamma < 1)$. We proceed as follows:

(a) we normalize $\underline{\theta}$ at one;

(b) we compute $q_{\tilde{i}}$ as $q_{\tilde{i}} = \min\{q \in \{q_1, \ldots, q_N\} : \gamma \leq \widehat{G}(q_i)\}$;

(c) when $q \in [q_1, q_{\tilde{i}})$, we know that $\gamma > \widehat{G}(q)$, so we can estimate $\nu'(q)$ as

$$\widehat{\nu'}(q) = \frac{T'(q)\left[\widehat{\gamma} - \widehat{G}(q_i)\right]^{1-\frac{\widehat{c}}{T'(q)}}}{\widehat{\gamma}^{1-\frac{c}{T'(q_1)}}} \exp\left\{-\widehat{c}\sum_{j=1}^{i-1} \log\left[\widehat{\gamma} - \widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right.$$
$$\left. -\widehat{c}\log\left[\widehat{\gamma} - \widehat{G}(q_i)\right]\left[\frac{1}{T'(q_i)} - \frac{1}{T'(q)}\right]\right\} \tag{64}$$

for $q \in [q_i, q_{i+1})$, $i = 1, \ldots, \widetilde{i} - 1$;

(d) from $\lim_{q \uparrow q_{\tilde{i}}}[\widehat{\gamma} - \widehat{G}(q)]^{1-\frac{\widehat{c}}{T'(q)}} = 1$, at $q_{\tilde{i}}$ we compute

$$\lim_{q \uparrow q_{\tilde{i}}} \widehat{\nu'}(q) = \frac{T'(q_{\tilde{i}})}{\widehat{\gamma}^{1-\frac{c}{T'(q_1)}}} \exp\left\{-\widehat{c}\sum_{j=1}^{\tilde{i}-1} \log\left[\widehat{\gamma} - \widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}.$$

(Even if there is no quantity at which $\widehat{\gamma} = \widehat{G}(q_{\tilde{i}})$, which means that no issue of singularity arises, the problem still exists that, by construction, $\gamma < \widehat{G}(q_{\tilde{i}})$ so $\log[\widehat{\gamma} - \widehat{G}(q_{\tilde{i}})]$ would be ill-defined. Computing $\nu'(q_{\tilde{i}})$ as a left-limit as $q \uparrow \widetilde{q}_i$ addresses this issue and allows us to derive an estimate for $\widetilde{\theta}_i$, which we can use to estimate $\nu'(q)$ at quantities above $\widetilde{q}_i$. See below for details.) Using the fact that $\theta_{\widetilde{i}}\nu'(q_{\widetilde{i}}) = T'(q_{\widetilde{i}})$, we obtain

$$\widehat{\theta}_{\widetilde{i}} = \widehat{\gamma}^{1-\frac{\widehat{c}}{T'(q_1)}} \exp\left\{\widehat{c}\sum_{j=1}^{\widetilde{i}-1} \log\left[\widehat{\gamma} - \widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}.$$

For accuracy, we estimate instead

$$\lim_{q\uparrow q_{\widetilde{i}}} \widehat{\nu'}(q) = \frac{T'(q_{\widetilde{i}})\left[\widehat{\gamma} - \widehat{G}(q_{\widetilde{i}-1})\right]^{1-\frac{\widehat{c}}{T'(q_{\widetilde{i}})}}}{\widehat{\gamma}^{1-\frac{c}{T'(q_1)}}} \exp\left\{-\widehat{c}\sum_{j=1}^{\widetilde{i}-1} \log\left[\widehat{\gamma} - \widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}$$

and, from $\widehat{\theta}_{\widetilde{i}} = T'(q_{\widetilde{i}})/\lim_{q\uparrow q_{\widetilde{i}}} \widehat{\nu'}(q)$, we obtain

$$\widehat{\theta}_{\widetilde{i}} = \widehat{\gamma}^{1-\frac{c}{T'(q_1)}} \left[\widehat{\gamma} - \widehat{G}(q_{\widetilde{i}-1})\right]^{\frac{\widehat{c}}{T'(q_{\widetilde{i}})}-1} \exp\left\{\widehat{c}\sum_{j=1}^{\widetilde{i}-1} \log\left[\widehat{\gamma} - \widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\}$$

whereas, from $\widehat{\theta}_{\widetilde{i}-1} = T'(q_{\widetilde{i}-1})/\widehat{\nu'}(q_{\widetilde{i}-1})$, we obtain

$$\widehat{\theta}_{\widetilde{i}-1} = \widehat{\gamma}^{1-\frac{\widehat{c}}{T'(q_1)}} \left[\widehat{\gamma} - \widehat{G}(q_{\widetilde{i}-1})\right]^{\frac{\widehat{c}}{T'(q_{\widetilde{i}-1})}-1} \exp\left\{\widehat{c}\sum_{j=1}^{\widetilde{i}-2} \log\left[\widehat{\gamma} - \widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\};$$

(e) when $q \in [q_{\widetilde{i}}, q_N)$, we compute

$$\widehat{\nu'}(q) = \frac{T'(q)[\widehat{G}(q_i) - \widehat{\gamma}]^{1-\frac{\widehat{c}}{T'(q)}}}{\widehat{\theta}_{\widetilde{i}}[\widehat{G}(q_{\widetilde{i}}) - \widehat{\gamma}]^{1-\frac{\widehat{c}}{T'(\widetilde{q}_i)}}} \exp\left\{-\widehat{c}\sum_{j=\widetilde{i}}^{i-1} \log\left[\widehat{G}(q_j) - \widehat{\gamma}\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right.$$
$$\left. -\widehat{c}\log\left[\widehat{G}(q_i) - \widehat{\gamma}\right]\left[\frac{1}{T'(q_i)} - \frac{1}{T'(q)}\right]\right\}$$

with $q \in [q_i, q_{i+1})$, $i = \widetilde{i}, \ldots, N-1$ (ignoring the knife-edged case in which $\gamma = \widehat{G}(q_{\widetilde{i}})$: if this case arises, perform this calculation at $i = \widetilde{i}+1, \ldots, N-1$). At q_N, we compute

$$\widehat{\nu'}(q_N) = \frac{T'(q_N)[\widehat{G}(q_N) - \widehat{\gamma}]^{1-\frac{\widehat{c}}{T'(q_N)}}}{\widehat{\theta}_{\widetilde{i}}[\widehat{G}(q_{\widetilde{i}}) - \widehat{\gamma}]^{1-\frac{\widehat{c}}{T'(\widetilde{q}_i)}}} \exp\left\{-\widehat{c}\sum_{j=\widetilde{i}}^{N-1} \log\left[\widehat{G}(q_j) - \widehat{\gamma}\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right\};$$

(f) we compute $\widehat{\theta}_i$, $i = 1, \ldots, N-1$, as $\widehat{\theta}_i = T'(q_i)/\widehat{\nu'}(q_i)$.

Case 2 ($\gamma = 0$). Note that in this case $q_{\widetilde{i}} = q_1$. Recall that $\widehat{G}(q_1) > 0$. We proceed as under (e)-(f) in Case 1. Since in this case $\widehat{\theta}_{\widetilde{i}} = \theta_1 = 1$, we obtain

$$\widehat{\nu'}(q) = \frac{T'(q)[\widehat{G}(q_i)]^{1-\frac{\widehat{c}}{T'(q)}}}{\theta_1[\widehat{G}(q_1)]^{1-\frac{\widehat{c}}{T'(q_1)}}} \exp\left\{-\widehat{c}\sum_{j=1}^{i-1} \log\left[\widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right.$$

$$-\widehat{c}\log\left[\widehat{G}(q_i)\right]\left[\frac{1}{T'(q_i)}-\frac{1}{T'(q)}\right]\Bigg\}$$

with $q \in [q_i, q_{i+1})$, $i = 1, \ldots, N-1$, and

$$\widehat{\nu}'(q_N) = \frac{T'(q_N)\widehat{G}(q_N)^{1-\frac{\widehat{c}}{T'(q_N)}}}{\theta_1[\widehat{G}(q_1)]^{1-\frac{\widehat{c}}{T'(q_1)}}}\exp\left\{-\widehat{c}\sum_{j=1}^{N-1}\log\left[\widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)}-\frac{1}{T'(q_{j+1})}\right]\right\}.$$

Case 3 ($\gamma = 1$). We proceed as in the case of the standard nonlinear pricing model. □

To compute consumer and producer surplus, and to perform the counterfactual experiment in which the seller is restricted to linear prices, we proceed as detalied above in Subsection C.1.1 Step 7 and in Subsection C.1.2.

C.2.2 Weakly-Convex Case

Suppose the weakly-convex case applies. We proceed according to the following steps.

Step 1. Compute $\widehat{G}(q_i)$, $T'(q_i)$, and $T''(q_i)$, $i = 1, \ldots, N$ from data on quantity purchases and prices (unit values) in each village. In particular, we compute

$$\widehat{G}(q) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}(q_i \leq q)$$

where for $q \in [\underline{q}, \overline{q}]$ and $\mathbf{1}(\cdot)$ is an indicator function.

Step 2. Compute $\widehat{c} = T'(q_1)$ (it should be the same as $\widehat{c} = T'(q_N)$) and q_{i_1} as

$$q_{i_1} = \max\{q_i \in \{q_1, \ldots, q_N\} : \sum_{j=1}^{i}[T'(q_j) - c]/T'(q_j)\Pr(q_j) \leq 0\}.$$

Step 3. Normalize $\underline{\theta} = \theta_1$ to one.

Step 4. For $q \in [\underline{q}, q_{i_1})$ where $\gamma = 0$, proceed analogously to the highly-convex case (Case 2):

- compute $\widehat{\nu}'(q_1) = T'(q_1)$ and from this obtain $\widehat{\theta}_{i_1}$ as

$$\widehat{\theta}_{i_1} = \frac{\widehat{G}(q_1)^{1-\frac{\widehat{c}}{T'(q_1)}}}{\widehat{G}(q_{i_1})^{1-\frac{\widehat{c}}{T'(q_{i_1})}}}\exp\left\{\widehat{c}\sum_{j=1}^{i_1-1}\log\left[\widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)}-\frac{1}{T'(q_{j+1})}\right]\right\}$$

since, by definition,

$$\widehat{\nu}'(q_1) = \frac{T'(q_1)\widehat{G}(q_1)^{1-\frac{\widehat{c}}{T'(q_1)}}}{\widehat{\theta}_{i_1}\widehat{G}(q_{i_1})^{1-\frac{\widehat{c}}{T'(q_{i_1})}}}\exp\left\{\widehat{c}\sum_{j=1}^{i_1-1}\log\left[\widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)}-\frac{1}{T'(q_{j+1})}\right]\right\};$$

- at $q \in [q_i, q_{i+1})$, $i = 1, \ldots, i_1 - 1$, compute

$$\widehat{\nu}'(q) = \frac{T'(q)\widehat{G}(q_i)^{1-\frac{\widehat{c}}{T'(q)}}}{\widehat{\theta}_{i_1}\widehat{G}(q_{i_1})^{1-\frac{\widehat{c}}{T'(q_{i_1})}}} \exp\left\{\widehat{c}\sum_{j=i+1}^{i_1-1} \log\left[\widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right.$$

$$\left. + \widehat{c}\log\left[\widehat{G}(q_i)\right]\left[\frac{1}{T'(q)} - \frac{1}{T'(q_{i+1})}\right]\right\};$$

- compute $\widehat{\theta}_i = T'(q_i)/\widehat{\nu}'(q_i)$, $i = 1, \ldots, i_1 - 1$.

Step 5. For $q \in [q_{i_1}, q_{\tilde{i}})$ where $\gamma(\theta(q)) = [G(q) - G(q_{i_1})]/[G(q_{i_2}) - G(q_{i_1})] < G(q)$:

- compute (note the difference compared to the highly-convex case: here the pontential point of singularity for $\gamma(\theta(q)) \leq G(q)$ is approached from below)

$$q_{\tilde{i}} = \max\{q_i \in \{q_1, \ldots, q_N\} : \gamma(\theta(q)) \leq G(q)\};$$

- compute

$$q_{i_2} = \min_{q_i \in \{q_1, \ldots, q_N\}} |G(q_i) - \left\{1 - G(q_{i_1})\left[\frac{1}{G(q_{\tilde{i}})} - 1\right]\right\}|;$$

- compute

$$\widehat{G}(q) - \widehat{\gamma}(\theta(q)) = \frac{\widehat{G}(q)\left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1\right] + \widehat{G}(q_{i_1})}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1})};$$

- at $q \in [q_i, q_{i+1})$, $i = i_1, \ldots, \widetilde{i} - 1$, compute

$$\widehat{\nu}'(q) = \frac{T'(q)\left\{\widehat{G}(q_i)\left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1\right] + \widehat{G}(q_{i_1})\right\}^{\frac{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1})}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1}\left[1 - \frac{c}{T'(q)}\right]}}{\widehat{\theta}_{i_1}\left\{\widehat{G}(q_{i_1})\left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1\right] + \widehat{G}(q_{i_1})\right\}^{\frac{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1})}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1}\left[1 - \frac{c}{T'(q_{i_1})}\right]}}$$

$$\cdot \exp\left(-\frac{\widehat{c}\left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1})\right]}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1}\sum_{j=i_1}^{i-1} \log\left\{\widehat{G}(q_j)\left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1\right] + \widehat{G}(q_{i_1})\right\}\left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})}\right]\right.$$

$$\left. - \frac{\widehat{c}\left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1})\right]}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1}\log\left\{\widehat{G}(q_i)\left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1\right] + \widehat{G}(q_{i_1})\right\}\left[\frac{1}{T'(q_i)} - \frac{1}{T'(q)}\right]\right)$$

and $\widehat{\theta}_i = T'(q_i)/\widehat{\nu}'(q_i)$;

- at $q_{\widetilde{i}}$: when $\gamma(\theta(q_{\widetilde{i}})) < G(q_{\widetilde{i}})$, we compute

$$\widehat{\nu}'(q_{\widetilde{i}}) = \frac{T'(q_{\widetilde{i}}) \left\{ \widehat{G}(q_{\widetilde{i}}) \left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1 \right] + \widehat{G}(q_{i_1}) \right\}^{\frac{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1})}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1} \left[1 - \frac{c}{T'(q_{\widetilde{i}})} \right]}}{\widehat{\theta}_{i_1} \left\{ \widehat{G}(q_{i_1}) \left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1 \right] + \widehat{G}(q_{i_1}) \right\}^{\frac{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1})}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1} \left[1 - \frac{c}{T'(q_{i_1})} \right]}}$$

$$\cdot \exp \left(- \frac{\widehat{c} \left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) \right]}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1} \sum_{j=i_1}^{\widetilde{i}-1} \log \left\{ \widehat{G}(q_j) \left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1 \right] + \widehat{G}(q_{i_1}) \right\} \left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})} \right] \right)$$

and $\widehat{\theta}_{\widetilde{i}} = T'(q_{\widetilde{i}}) / \widehat{\nu}'(q_{\widetilde{i}})$;

- at $q_{\widetilde{i}}$: when $\gamma(\theta(q_{\widetilde{i}})) = G(q_{\widetilde{i}})$, we compute $\widehat{\nu}'(q_{\widetilde{i}})$ as

$$\lim_{q \uparrow q_{\widetilde{i}}} \widehat{\nu}'(q) = \frac{T'(q_{\widetilde{i}}) \left\{ \widehat{G}(q_{\widetilde{i}-1}) \left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1 \right] + \widehat{G}(q_{i_1}) \right\}^{\frac{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1})}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1} \left[1 - \frac{c}{T'(q_{\widetilde{i}})} \right]}}{\widehat{\theta}_{i_1} \left\{ \widehat{G}(q_{i_1}) \left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1 \right] + \widehat{G}(q_{i_1}) \right\}^{\frac{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1})}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1} \left[1 - \frac{c}{T'(q_{i_1})} \right]}}$$

$$\cdot \exp \left(- \frac{\widehat{c} \left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) \right]}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1} \sum_{j=i_1}^{\widetilde{i}-1} \log \left\{ \widehat{G}(q_j) \left[\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1}) - 1 \right] + \widehat{G}(q_{i_1}) \right\} \left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})} \right] \right)$$

and $\widehat{\theta}_{\widetilde{i}} = T'(q_{\widetilde{i}}) / \lim_{q \uparrow \widetilde{q}_i} \widehat{\nu}'(q)$.

Step 6. For $q \in [q_{\widetilde{i}}, q_{i_2})$ where $\gamma(\theta(q)) = [G(q) - G(q_{i_1})]/[G(q_{i_2}) - G(q_{i_1})] > G(q)$:

- compute the difference $\widehat{\gamma}(\theta(q)) - \widehat{G}(q)$ as

$$\widehat{\gamma}(\theta(q)) - \widehat{G}(q) = \frac{\widehat{G}(q) \left[1 - \widehat{G}(q_{i_2}) + \widehat{G}(q_{i_1}) \right] - \widehat{G}(q_{i_1})}{\widehat{G}(q_{i_2}) - \widehat{G}(q_{i_1})};$$

- at $q \in [q_i, q_{i+1})$, $i = \widetilde{i}, \ldots, i_2 - 1$ (from Step 5: if $q_{\widetilde{i}}$ is truly a point of singularity, then the relevant intervals are indexed by $i = \widetilde{i} + 1, \ldots, i_2 - 1$), compute marginal utility as

$$\widehat{\nu}'(q) = \frac{T'(q) \left\{ [1 - G(q_{i_2}) + G(q_{i_1})] G(q_{\widetilde{i}}) - G(q_{i_1}) \right\}^{\frac{G(q_{i_2}) - G(q_{i_1})}{1 - G(q_{i_2}) + G(q_{i_1})} \left[1 - \frac{c}{T'(q_{\widetilde{i}})} \right]}}{\widehat{\theta}_{\widetilde{i}} \left\{ [1 - G(q_{i_2}) + G(q_{i_1})] G(q_i) - G(q_{i_1}) \right\}^{\frac{G(q_{i_2}) - G(q_{i_1})}{1 - G(q_{i_2}) + G(q_{i_1})} \left[1 - \frac{c}{T'(q)} \right]}}$$

$$\cdot \exp \left\{ \frac{c[G(q_{i_2}) - G(q_{i_1})]}{1 - G(q_{i_2}) + G(q_{i_1})} \sum_{j=\widetilde{i}}^{i-1} \log \left\{ [1 - G(q_{i_2}) + G(q_{i_1})] G(q_j) - G(q_{i_1}) \right\} \left[\frac{1}{T'(q_j)} - \frac{1}{T'(q_{j+1})} \right] \right.$$
$$\left. + \frac{c[G(q_{i_2}) - G(q_{i_1})]}{1 - G(q_{i_2}) + G(q_{i_1})} \log \left\{ [1 - G(q_{i_2}) + G(q_{i_1})] G(q_i) - G(q_{i_1}) \right\} \left[\frac{1}{T'(q_i)} - \frac{1}{T'(q)} \right] \right\}$$

and to obtain $\widehat{\nu}'(q)$ at q_{i_2}, compute

$$\widehat{\nu}'(q_{i_2}) = \frac{T'(q_{i_2})\left\{[1-G(q_{i_2})+G(q_{i_1})]G(q_{\widetilde{i}})-G(q_{i_1})\right\}^{\frac{G(q_{i_2})-G(q_{i_1})}{1-G(q_{i_2})+G(q_{i_1})}\left[1-\frac{c}{T'(q_{\widetilde{i}})}\right]}}{\widehat{\theta}_{\widetilde{i}}\left\{[1-G(q_{i_2})+G(q_{i_1})]G(q_{i_2})-G(q_{i_1})\right\}^{\frac{G(q_{i_2})-G(q_{i_1})}{1-G(q_{i_2})+G(q_{i_1})}\left[1-\frac{c}{T'(q_{i_2})}\right]}}$$

$$\cdot \exp\left\{\frac{c[G(q_{i_2})-G(q_{i_1})]}{1-G(q_{i_2})+G(q_{i_1})}\sum_{j=\widetilde{i}}^{i_2-1}\log\left\{[1-G(q_{i_2})+G(q_{i_1})]G(q_j)-G(q_{i_1})\right\}\left[\frac{1}{T'(q_j)}-\frac{1}{T'(q_{j+1})}\right]\right\};$$

- for $i=\widetilde{i},\ldots,i_2$, compute $\widehat{\theta}_i = T'(q_{\widetilde{i}})/\widehat{\nu}'(q_{\widetilde{i}})$.

Step 7. For $q \in [q_{i_2}, q_N)$ where $\gamma(\theta(q)) = 1 > G(q)$:

- compute marginal utility

$$\widehat{\nu}'(q) = \frac{T'(q)\left[1-\widehat{G}(q_i)\right]^{1-\frac{\widehat{c}}{T'(q)}}}{\widehat{\theta}_{i_2}\left[1-\widehat{G}(q_{i_2})\right]^{1-\frac{\widehat{c}}{T'(q_{i_2})}}}\exp\left\{-\widehat{c}\sum_{j=i_2}^{i-1}\log\left[1-\widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)}-\frac{1}{T'(q_{j+1})}\right]\right.$$
$$\left.-\widehat{c}\log\left[1-\widehat{G}(q_i)\right]\left[\frac{1}{T'(q_i)}-\frac{1}{T'(q)}\right]\right\}$$

where

$$\lim_{q\uparrow q_N}\nu'(q) = \frac{T'(q_N)\exp\left\{-\widehat{c}\sum_{j=i_2}^{N-1}\log\left[1-\widehat{G}(q_j)\right]\left[\frac{1}{T'(q_j)}-\frac{1}{T'(q_{j+1})}\right]\right\}}{\widehat{\theta}_{i_2}\left[1-\widehat{G}(q_{i_2})\right]^{1-\frac{\widehat{c}}{T'(q_{i_2})}}};$$

- for $i=i_2,\ldots,N-1$, compute types as $\widehat{\theta}_i = T'(q_{\widetilde{i}})/\widehat{\nu}'(q_{\widetilde{i}})$ and $\widehat{\theta}_N = T'(q_N)/\lim_{q\uparrow q_N}\nu'(q)$.

To compute consumer and producer surplus, and to perform the counterfactual experiment in which the seller is restricted to linear prices, we proceed as detailed above in Subsection C.1.1 Step 7 and in Subsection C.1.2.

REFERENCES

ATTANASIO, O., V. DI MARO, V. LECHENE, and D. PHILLIPS (2009): "The Welfare Consequences of Increases in Food Prices in Rural Mexico and Colombia", mimeo.

ATTANASIO, O., and C. FRAYNE (2006): "Do the Poor Pay More?", mimeo.

BARON, D.P., and R.B. MYERSON (1982): "Regulating a Firm with Unknown Cost", *Econometrica* 50(4), 911-930.

GUERRE, E., I. PERRIGNE, and Q. VUONG (2000): "Optimal Nonparametric Estimation of First-Price Auctions", *Econometrica* 68(3), 525-574.

JULLIEN, B. (2000): "Participation Constraints in Adverse Selection Models", *Journal of Economic Theory* 93(1), 1-47.

LESLIE, P. (2004): "Price Discrimination in Broadway Theatre", *Rand Journal of Economics*

35 (3), 520-541.
MAS-COLELL, A., M.D WHINSTON, and J.R. GREEN (1995): "*Microeconomic Theory*", Oxford University Press, USA.
MASKIN, E., and J. RILEY (1984): "Monopoly with Incomplete Information", *Rand Journal of Economics* 15(2), 171-196.
MULLER, A., and D. STOYAN (2002): *Comparison Methods for Stochastic Models and Risks*, Wiley Series in Probability and Statistics, UK.
PERRIGNE, I., and Q. VUONG (2008): "Nonlinear Pricing in Yellow Pages", mimeo.
STOLE, L.A. (2006): "Price Discrimination and Competition", mimeo.